University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

1-1-2015

Putting Your Past Behind You: Why and How Fresh Starts Motivate Goal Pursuit--and When They Backfire

Hengchen Dai
University of Pennsylvania, daihengchenpku@gmail.com

Follow this and additional works at: http://repository.upenn.edu/edissertations

Part of the Organizational Behavior and Theory Commons, and the Social and Behavioral Sciences Commons

Putting Your Past Behind You: Why and How Fresh Starts Motivate Goal Pursuit--and When They Backfire

Abstract

People often fail to exercise the self-control required to tackle their goals and improve their performance. However, many feel that we have opportunities throughout our lives to start fresh with a clean slate (e.g., around New Year's Day, after a Catholic confession). Although the notion of "fresh starts" has long been endorsed by our culture, researchers have not systematically explored the implications of fresh starts for people's motivation to exert effort in goal-directed activities. Across three chapters, I examine (a) how and why fresh starts affect individuals' ability to exert the self-control needed to achieve their aspirations and (b) when fresh starts may adversely influence individuals' motivation to improve their performance.

In Chapter 1, three archival field studies demonstrate that people engage in aspirational behaviors (e.g., exercising, creating a goal commitment contract) more frequently at the start of new time periods that are initiated by temporal landmarks (e.g., the beginning of a new week/month/year/school semester, or immediately following a holiday, a school break, or a birthday). In Chapter 2, five laboratory studies show that meaningful temporal landmarks--dates imbued with meaning due to their identity-relevance or rarity--are more likely to spur goal pursuit than (a) ordinary days or (b) objectively identical but psychologically less meaningful landmarks. Further, I provide evidence for one mechanism underlying these findings: temporal landmarks (particularly meaningful landmarks) relegate past imperfections to a previous period, making the current self feel more capable of pursuing aspirations. Chapter 3 investigates the impact on individuals' future performance of tracking their performance without incorporating records of their past performance (a phenomenon I refer to as a performance reset). I propose that when individuals believe their past performance was poor, a performance reset will improve their performance by boosting self-efficacy and commitment. However, I expect resets to hurt performance by decreasing commitment without increasing self-efficacy when individuals believe their past performance was strong. One archival field study and four laboratory experiments support this hypothesized relationship between performance resets, past performance, and future performance; these studies also provide preliminary evidence that self-efficacy mediates this relationship. Chapter 4 discusses directions for future research.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Operations & Information Management

First Advisor
Katherine L. Milkman

PUTTING YOUR PAST BEHIND YOU: WHY AND HOW FRESH STARTS

MOTIVATE GOAL PURSUIT—AND WHEN THEY BACKFIRE

Hengchen Dai

A DISSERTATION

in

Operations and Information Management

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

Signature _____

Katherine L. Milkman, Assistant Professor of Operations and Information Management

Graduate Group Chairperson

Signature _____

Eric T. Bradlow, Professor of Marketing, Statistics, and Education

Dissertation Committee

Maurice E. Schweitzer, Professor of Operations and Information Management

Katherine Klein, Professor of Management

David A. Hofmann, Professor of Organizational Behavior/Strategy, University of North Carolina-Chapel Hill

DEDICATION

This dissertation is dedicated to my loving parents:

Bendong Dai and Zhengling Hu

ACKNOWLEDGMENT

ABSTRACT

PUTTING YOUR PAST BEHIND YOU: WHY AND HOW FRESH STARTS

MOTIVATE GOAL PURSUIT—AND WHEN THEY BACKFIRE

Hengchen Dai

Katherine L. Milkman

People often fail to exercise the self-control required to tackle their goals and improve their performance. However, many feel that we have opportunities throughout our lives to start fresh with a clean slate (e.g., around New Year's Day, after a Catholic confession). Although the notion of "fresh starts" has long been endorsed by our culture, researchers have not systematically explored the implications of fresh starts for people's motivation to exert effort in goal-directed activities. Across three chapters, I examine (a) how and why fresh starts affect individuals' ability to exert the self-control needed to achieve their aspirations and (b) when fresh starts may adversely influence individuals' motivation to improve their performance.

In Chapter 1, three archival field studies demonstrate that people engage in aspirational behaviors (e.g., exercising, creating a goal commitment contract) more frequently at the start of new time periods that are initiated by temporal landmarks (e.g., the beginning of a new week/month/year/school semester, or immediately following a holiday, a school break, or a birthday). In Chapter 2, five laboratory studies show that meaningful temporal landmarks—dates imbued with meaning due to their identity-relevance or rarity—are more likely to spur goal pursuit than (a) ordinary days or (b) objectively identical but psychologically less meaningful landmarks. Further, I provide

evidence for one mechanism underlying these findings: temporal landmarks (particularly meaningful landmarks) relegate past imperfections to a previous period, making the current self feel more capable of pursuing aspirations. Chapter 3 investigates the impact on individuals' future performance of tracking their performance without incorporating records of their past performance (a phenomenon I refer to as a *performance reset*). I propose that when individuals believe their past performance was poor, a performance reset will improve their performance by boosting self-efficacy and commitment. However, I expect resets to hurt performance by decreasing commitment without increasing self-efficacy when individuals believe their past performance was strong. One archival field study and four laboratory experiments support this hypothesized relationship between performance resets, past performance, and future performance; these studies also provide preliminary evidence that self-efficacy mediates this relationship. Chapter 4 discusses directions for future research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Although people often fail to exercise the self-control required to tackle their goals and improve their performance, many feel that they have opportunities throughout their lives to start fresh with a clean slate. For example, New Year's Day traditionally motivates people to exert greater than usual effort towards achieving their goals; religious purification ceremonies are deemed as opportunities for people to put their imperfections behind them. Despite that the notion of such "fresh starts" has long been endorsed by many cultures, researchers have not systematically explored the implications of fresh starts for people's motivation to engage in goal-directed activities and improve their performance.

This dissertation is dedicated to advance our understanding of how fresh starts affect motivation and individual performance. I define fresh starts broadly as moments or events that make people feel psychologically more distant from the past than they otherwise would. Fresh starts are pervasive in our lives. Transitions that signify the beginning of a new chapter in our lives (e.g., graduating from college; starting a first job; moving to a new city) may cause us to feel separated from who we were in the past (Bartels & Rips, 2010; Cantor, Norem, Niedenthal, Langston, & Brower, 1987). We may also feel further away from our past selves following certain personal and calendar events (e.g., birthdays, wedding anniversaries, Mondays, holidays) that do not change our environments but simply demarcate adjacent time periods in our lives (Peetz & Wilson, 2013, 2014). In addition, fresh starts may arise and separate us from our past success or failures when we have a clean performance record at the beginning of a new performance

period (e.g., at the outset of a sales season for salespeople, at the beginning of a semester for students). Understanding the effects of fresh starts on people's motivation to pursue their aspirations and improve their performance has the potential to deepen our understanding of what spurs goal pursuit and the exertion of self-control.

Across three chapters, I examine two types of fresh starts that are prevalent inside and outside of the workplace. The first type of fresh start includes naturally-arising occasions and calendar events (e.g., birthdays, holidays, the beginning of a work week) that demarcate the passage of time. The second type of fresh start occurs when individuals' performance is tracked from a new reference point and past performance is not separated from performance appraisals in the current performance period. I investigate (1) how and why fresh starts affect individuals' ability to exert the self-control needed to achieve their aspirations and (b) when fresh starts may adversely influence individuals' motivation to improve their performance.

<u>Overview of Chapter 1</u>

In Chapter 1 (co-authored with Katherine L. Milkman and Jason Riis), I investigate whether people feel particularly motivated to tackle their goals at certain naturally arising points in time. Specifically, I test the hypothesis that temporal landmarks—distinctive calendar dates (e.g., the beginning of the week/month) or notable life events (e.g., a birthday) that stand out from more ordinary dates in our lives—can spur goal pursuit. Across three field studies, I analyzed the frequency with which (a) Americans search for the term "diet" on Google, (b) undergraduate students visit a university gym, and (c) Internet users commit to pursuing a personal goal on a goal-

setting website. I document the existence of a "fresh start effect," whereby people engage in goal-directed, aspirational behaviors (i.e., dieting, exercising, and creating goal-commitment contracts) more frequently at the start of new periods demarcated by temporal landmarks, including the beginning of a new week, month, year, and school semester, as well as right after a public holiday, a school break, or a birthday.

Chapter 1 provides the first systematic investigation of naturally arising points in time when people are motivated to pursue their aspirations. The findings presented in this chapter add to a growing body of research demonstrating the impact of ubiquitous environmental factors and organizational cues on individuals' capacity to exert self-control and highlight the inaccuracy of models assuming that self-control is a time-invariant trait.

<div align="center">Overview of Chapter 2</div>

Chapter 2 (co-authored with Katherine L. Milkman and Jason Riis) builds on and extends Chapter 1 in three ways. First, Chapter 1 relies on archrival data analyses and provides *correlational* evidence that goal-related activities are more frequent following temporal landmarks. In Chapter 2, I experimentally manipulate the salience of temporal landmarks and test whether temporal landmarks *causally* spur goal pursuit. Second, I examine which types of temporal landmarks are especially likely to motivate goal-directed activities in Chapter 2. Across five laboratory studies, I demonstrate that meaningful temporal landmarks—dates imbued with meaning due to their identity-relevance or rarity—are more likely to spur goal pursuit than (a) ordinary days and (b) objectively identical but psychologically less meaningful landmarks. Further, I provide a

theoretical account for these findings, building on (a) the theory of temporal self-appraisal and (b) the notion that subjective perceptions of time are malleable. Specifically, I propose and show that temporal landmarks (particularly more meaningful ones) help an individual relegate past imperfections to her past self, making her current self feel superior and more capable of pursuing aspirations.

Chapter 2 brings together past psychological research on subjective perceptions of time and temporal self-appraisals with the literature on motivation. Further, it highlights that the non-linear way temporal landmarks cause us to perceive time can alter our motivation to pursue our goals. It also demonstrates that more meaningful temporal landmarks can create a greater separation between our temporal selves than less meaningful landmarks, encouraging us to strive for our goals more determinedly.

<p align="center">Overview of Chapter 3</p>

Taken together, Chapters 1 and 2 demonstrate that naturally arising points in time that initiate "new epochs" (e.g., the beginning of the week, month, season, or year) motivate people to exert efforts toward their self-improvement goals. In Chapter 3, I examine the motivating effects of fresh starts as well as when fresh starts may hamper individuals' motivation to expend effort, thus decreasing their overall performance. In addition, I explore a different type of fresh start than examined in Chapters 1 and 2: I explore the type of fresh start associated with the outset of a new performance tracking period in an organization.

In many organizations, metrics tracking performance are reset to zero at the start of a new performance period. For example, for a salesperson, monthly sales totals might

be reset at the beginning of each month. Additionally, some milestone events may cause organizations to disregard individuals' historic performance records (e.g., a relocation to a different branch office). Chapter 3 examines the effects on individuals' future performance of "performance resets" or tracking individuals' performance from a new starting point after excluding their past performance record..

Drawing on the literature about performance feedback, self-efficacy, commitment, and motivation, I propose that the effects of a performance reset on future performance will depend on an individual's past performance. I argue that this can be explained by changes in an individual's self-efficacy and commitment to performing well. Specifically, I expect that when individuals perceive their past performance as poor, resetting their performance statistics should (a) boost their self-efficacy because they now view past failures as less indicative of their future ability and (b) increase their commitment to working hard because they now have a new opportunity to establish their capability. However, when individuals believe their past performance was strong, resetting their record should (a) have no benefits for self-efficacy because they will already have a positive self-view and (b) decrease commitment because once past achievements are set in stone, the value of proving themselves is reduced. In light of the well-known positive relationship between self-efficacy, commitment, and future performance, I predict that performance resets will improve future performance when individuals performed well in the past but dampen future performance when individuals performed poorly.

Across an archival field study and four laboratory experiments, I find that performance resets improve performance when individuals view their past performance as poor, but reset harm performance when individuals view their past performance as strong consistent with my hypothesis. I provide preliminary evidence that self-efficacy mediates the relationship between performance resets, past performance, and future performance, but I find no support for goal commitment as a mediator.

Research on motivation and performance management has previously shown that an individual's past performance significantly influences her self-efficacy, commitment, and future performance (e.g., Bandura, 1991, 1997; Klein & Kim, 1998; Locke, Latham, & Erez, 1988). However, there has been sparse research on how an individual's psychological separation from her past (poor or strong) performance may affect her motivation and effort allocation moving forward. Chapter 3 contributes to the management literature by investigating conditions under which the experience of having a clean slate may increase versus decrease an individual's motivation and performance.

<div align="center">Summary</div>

This dissertation examines fresh starts induced by two types of events: temporal landmarks and performance resets. It provides a theoretical framework that predicts when a fresh start will inspire an individual to improve her performance (working harder to achieve personal goals or tasks at work) as well as when a fresh start may be demotivating. Chapter 4 discusses directions for future research.

In addition to contributing to the psychology and management literatures, this dissertation offers practical insights into how managers can help employees pursue their

personal goals and improve their productivity. For example, messages and interventions designed to promote goal-directed activities may be more effective when provided after landmark events at work (e.g., the completion of a project, a change in the company's leadership) and particularly after work events that feel especially meaningful (e.g., an employee's 10th work anniversary, the first time in years an employee has moved to a new office). Further, when deciding whether and when to reset employees' performance metrics to zero, managers should consider that performance resets may affect employees differently depending on their past performance.

References

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287. doi: 10.1016/0749-5978(91)90022-L

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.

Bartels, D. M., & Rips, L. J. (2010). Psychological connectedness and intertemporal choice. *Journal of experimental psychology. General, 139*(1), 49-69. doi: 10.1037/a0018062

Cantor, N., Norem, J. K., Niedenthal, P. M., Langston, C. A., & Brower, A. M. (1987). Life tasks, self-concept ideals, and cognitive strategies in a life transition. *Journal of Personality and Social Psychology, 53*(6), 1178-1191. doi: 10.1037/0022-3514.53.6.1178

Klein, H. J., & Kim, J. S. (1998). A field study of the influence of situational constraints, leader-member exchange, and goal commitment on performance. *Academy of Management Journal, 41*(1), 88-95. doi: 10.2307/256900

Locke, E. A., Latham, G. P., & Erez, M. (1988). The determinants of goal commitment. *Academy of Management Review, 13*(1), 23-39. doi: 10.2307/258352

Peetz, J., & Wilson, A. E. (2013). The post-birthday world: Consequences of temporal landmarks for temporal self-appraisal and motivation. *Journal of Personality and Social Psychology, 104*(2), 249-266. doi: 10.1037/A0030477

Peetz, J., & Wilson, A. E. (2014). Marking time: Selective use of temporal landmarks as

barriers between current and future selves. *Personality and Social Psychology

Bulletin, 40*(1), 44-56. doi: 10.1177/0146167213501559

CHAPTER 1

THE FRESH START EFFECT:

TEMPORAL LANDMARKS MOTIVATE ASPIRATIONAL BEHAVIOR

Hengchen Dai, Katherine L. Milkman and Jason Riis

ABSTRACT

The popularity of New Year's resolutions suggests that people are more likely to tackle their goals immediately following salient temporal landmarks. If true, this little-researched phenomenon has the potential to help people overcome important willpower problems that often limit goal attainment. Across three archival field studies, we provide evidence of a "fresh start effect." We show that Google searches for the term "diet" (Study 1), gym visits (Study 2), and commitments to pursue goals (Study 3) all increase following temporal landmarks (e.g., the outset of a new week, month, year, or semester; a birthday; a holiday). We propose that these landmarks demarcate the passage of time, creating many new mental accounting periods each year, which relegate past imperfections to a previous period, induce people to take a big-picture view of their lives, and thus motivate aspirational behaviors.

Keywords: goals; motivation; self-control; temporal landmarks; mental accounting

Introduction

The beginning of the year is widely documented as a time when millions of people commit themselves with atypical vigor to achieving their goals, such as losing weight, eating more healthfully, quitting smoking, obtaining a better education, and saving more money (Marlatt & Kaplan, 1972; Norcross, Mrykalo, & Blagys, 2002). The U.S. government actually lists popular New Year's resolutions on its official website and provides resources to help its citizens tackle their goals in the coming year (USA.gov, 2013). More broadly, the notion that fresh starts are possible and offer individuals an opportunity to improve themselves has long been endorsed by our culture. For example, Christians can be "born again," Catholic confessions and penance provide sinners with a fresh start, many religious groups engage in ritual purification or ablution ceremonies (e.g., Buddhists, Christians, Muslims, and Jews), and the metaphorical phoenix rising from the ashes is a ubiquitous symbol of rebirth. This suggests a widely shared belief that we have opportunities throughout our lives to start fresh with a clean slate, with the "New Year's effect" representing just one example of a far broader phenomenon documented in this paper. Specifically, we show that special occasions and calendar events (e.g., a birthday, a holiday, the beginning of a new week/month), which demarcate the passage of time and create numerous "fresh start" opportunities at the beginning of new cycles throughout each year, are associated with subsequent increases in aspirational behavior.

Understanding when people are most motivated to pursue their aspirations is important for a number of reasons. Aspirational behaviors are activities that help people achieve their wishes and personal goals (Merriam-Webster.com, 2013). Examples of

behaviors that people frequently aspire to engage in more often include exercising, saving money, studying, dating, and dieting (Khan, Dhar, & Wertenbroch, 2005). Notably, we often lack the self-control to expend the time and effort needed to achieve our aspirations and instead postpone the work necessary to tackle our goals until a later date (Bazerman, Tenbrunsel, &Wade-Benzoni, 1998; Milkman, Rogers, & Bazerman, 2008; O'Donoghue & Rabin, 1999). For example, individuals often repeatedly procrastinate when it comes to dieting, exercising, and quitting smoking. Over time, such near-sighted decision making can result in serious individual and societal problems, such as high rates of obesity and cancer.

Many researchers have sought to understand situational factors that motivate people to pursue their aspirations (e.g., Shiv & Fedorikhin, 1999; Botti et al., 2008; Sela, Berger, & Liu, 2009; Milkman, 2012; Toure-Tillery & Fishbach, 2012). However, sparse research has investigated naturally-arising points in time when people feel particularly motivated to tackle their goals. Notable exceptions include past work demonstrating increased attention to aspirations at the outset of the new year (Marlatt & Kaplan, 1972; Norcross et al., 2002) as well as unpublished (Cross, Peretz, Munoz-LaBoy, Lapp, Shelley, & Rosenfield, 2006; Fry & Neff, 2010) and concurrent studies (Ayers, Althouse, Johnson, & Cohen, 2014) suggesting that people are most likely to think about their health on Mondays.

This paper empirically examines whether other points in time, beyond (but including) the start of a new year or week, are associated with increases in aspirational behavior. Across three field studies, we demonstrate that people are more likely to pursue

various types of aspirational behavior (e.g., dieting, exercising, goal pursuit) at the start

of "new epochs" initiated by the incidence of temporal landmarks, including the

beginning of a new week, month, year, and school semester, as well as immediately

following a public holiday, a school break, or a birthday. We use historical Google search

volume data, university gym attendance records, and data from the goal-setting website

(www.stickK.com; hereafter referred to as stickK) to document this phenomenon, which

we call "the fresh start effect." Though much past research assumes that self-control is a

time-invariant trait (e.g., Shoda, Mischel, & Peake, 1990), we add to a growing body of

recent research suggesting that self-control capacity is variable (Shiv & Fedhorkin, 1999;

Khan & Dhar, 2006).

We postulate that temporal landmarks, including personally meaningful events

(e.g., birthdays, job changes) and socially constructed calendar partitions (e.g., the outset

of a new month, the observance of a public holiday), demarcate the passage of time and

open new mental accounting periods. We propose two primary explanations for the fresh

start effect. Specifically, we propose that naturally-arising time markers (a) create

discontinuities in time perceptions that make people feel disconnected from their past

imperfections; and (b) disrupt people's focus on day-to-day minutiae, thereby promoting

a big-picture view of life. We postulate that these processes triggered by fresh start

moments encourage people to pursue their aspirations. We will address and rule out a

number of key alternative explanations for our findings, but it is important to

acknowledge that our field data provide imperfect insights into the mechanisms

responsible for the fresh start effect and thus additional future research on this topic would be extremely valuable.

Conceptual Framework

Temporal Landmarks Segregate Life into Numerous, Distinct Mental Accounting Periods

Past research on mental accounting has demonstrated that "choices are altered by the introduction of notional…boundaries" (Thaler, 1999, p.197) and has largely focused on examining how the initiation of new mental accounting periods affects financial outcomes (for reviews, see Read, Loewenstein, & Rabin, 1999; Thaler, 1999; Soman, 2004; Soman & Ahn, 2011). While this previous research has shown that time is not treated as continuous and fungible (Rajagopal & Rha, 2009; Soman, 2001), many implications of the non-linear way in which we experience time have not yet been explored. In this paper, we investigate how people's motivation to pursue personal goals can be altered by the initiation of new mental accounting periods, as demarcated by temporal landmarks.

Temporal landmarks, or distinct events that "stand in marked contrast to the seemingly unending stream of trivial and ordinary occurrences that happen to us everyday" (Shum, 1998, p.423), have been shown to structure our memories and experiences (Robinson, 1986; Shum, 1998). One type of temporal landmark includes reference points on socially constructed and shared timetables. Examples include the beginning of an academic semester, secular and religious holidays, and time dividers on the yearly calendar (Kurbat, Shevell, & Rips, 1998; Robinson, 1986). Another type of

temporal landmark includes personally-relevant life events that demarcate our personal histories, such as developmental milestones, life transitions, first experiences, and occasions of recurrent significance (Robinson, 1986; Rubin & Kozin, 1984). These temporal landmarks not only influence the manner in which people recall memories, experiences, and time durations retrospectively (Ahn, Liu, & Soman, 2009; Rubin & Kozin, 1984; Shum, 1998; Zauberman, Levav, Diehl, & Bhargave, 2010), but they are also used to organize current activities and future plans and to designate the boundaries of temporal periods (LeBoeuf, Williams, & Brenner, 2014; Peetz & Wilson, 2013; Robinson, 1986; Soster, Monga, & Bearden, 2010; Tu & Soman, 2014). For example, when asked to describe the periods into which they divide their time, people frequently list cycles such as a *day*, *week*, *month*, *school semester*, and *school break* (Soster et al., 2010). Furthermore, when a salient temporal landmark (e.g., a public holiday, a birthday, a school event) in between two points in time is highlighted, people are more likely to perceive those two points in time as arising in two distinct periods (Peetz & Wilson, 2013; Soster et al. , 2010; Tu & Soman, 2014). Together, this research suggests that temporal landmarks open new mental accounts. We propose that when temporal landmarks open new mental accounts, the beginning of a new period stands in contrast to more typical days in our lives. Below, we describe two perspectives on why temporal landmarks may then motivate people to pursue their aspirations.

Temporal Landmarks Relegate Past Imperfections to a Previous Mental Accounting Period

Individuals think of their past, current, and future selves as interconnected but separable components of their identity (Parfit, 1984) and often compare these selves to one another (Wilson & Ross, 2001). For example, an individual might consider whether she is a wiser person now than she was in the past, or she might plan to be a better person in the future.

Past research has shown that the perceived connection between our present and past temporal selves can be affected by (a) personally-relevant events such as a religious conversion (Libby & Eibach, 2002, 2011; Wilson & Ross, 2003; Bartels & Rips, 2010) and (b) the salience of calendar landmarks (Peetz & Wilson, 2013). Anecdotally, past researchers have noted that people who change (e.g., receive a cancer diagnosis, recover from an addiction) often describe their pre-change self as a distinct person (Libby & Eibach, 2002). Wilson and Ross (2003) suggest that many real-life experiences, ranging from personal milestones (e.g., a marriage or job change) to mundane changes in appearance or possessions (e.g., getting a new haircut or suit) can distance us from our past self. Together, this research demonstrates that landmarks in people's lives generate a disassociation between present and past selves.[1]

We propose that the psychological separation between one's present and past selves induced by temporal landmarks motivates people to pursue their aspirations. The theory of temporal self-appraisal contends that people evaluate their past self in a manner

---

[1] Recent research has also shown that temporal landmarks affect the perceived psychological distance between people's present and *future* selves. Bartels and Rips (2010) demonstrated that the psychological connectedness between a person's present and future selves can be weakened by prompting them to imagine experiencing landmark events (e.g., finding out that they were adopted, being imprisoned as a political hostage). Also, recent work showed that highlighting a future landmark event (e.g., a public holiday, a birthday) induces a psychological separation between the current self and the post-landmark future self (Peetz & Wilson, 2013).

that flatters their current self (Wilson & Ross, 2001). In particular, people tend to disparage and attribute their past failures to their former, distant self because (a) faults of a remote, past self are less apt to tarnish their present self-image and thus are less threatening and (b) criticizing a distant, inferior self implies self-improvement over time, which is viewed as desirable (Wilson & Ross, 2001). Importantly, temporal landmarks – moments that psychologically disconnect one's past, current and future selves – lead people to perceive a contrast between their disconnected selves (Peetz & Wilson, 2013). This facilitates a tendency to view one's past self as inferior and one's current self as superior (Wilson & Ross, 2001).

We argue that by relegating previous imperfections to a past self and generating a sense that the current self is superior, temporal landmarks can alter people's decisions. Considerable past research has shown that people are motivated to maintain a coherent self-image (Epstein, 1973; Markus, Mullally, & Kitayama, 1997; Kivetz & Tyler, 2007). For example, if people perceive themselves as moral, they are more likely to pursue moral actions (Aquino & Reed, 2002). Thus, when people perceive themselves to be superior to a past self (e.g., more self-disciplined, more extroverted, etc.), past research suggests they will behave in accordance with those perceptions (e.g., study harder, become more active in social events, etc.). Therefore, we hypothesize that when temporal landmarks psychologically disconnect us from our inferior, past self and make us feel superior, we will be motivated to behave better than we have in the past and strive with enhanced fervor to achieve our aspirations.

It should be noted that some people may not see their past self as inferior to their current self. However, so long as the average person sees her past self as more flawed than her current self, the fresh start effect should emerge on average, albeit not necessarily for every individual.

Temporal Landmarks Promote a Focus on the Big Picture

In addition to psychologically separating people from their past imperfections, temporal landmarks may motivate people to pursue their aspirations by altering the manner in which they process information and form preferences. Specifically, by creating discontinuities in our perceptions of time, experiences, and activities, temporal landmarks may promote taking a broader view of decisions. Liu (2008) shows that interruptions to decision making (e.g., switching to a new background task while pondering a focal decision) change information processing. Specifically, interruptions move people from a bottom-up, contextually rich mode of thinking focused on concrete data to a higher level, top-down mode guided by pre-existing goal and knowledge structures. Temporal landmarks may serve as one type of disruption to decision making and thus direct attention to high-level, goal-relevant information. Indeed, there is some evidence that this is the case. For example, Bhargave and Miron-shatz (2012) show that people at milestone ages (e.g., 30, 40 years old) are more likely than those at other ages to judge their life satisfaction based on their overall achievements rather than their daily emotions, highlighting that temporal landmarks can lead to bigger picture thinking.

Past research has shown that high-level, big picture thinking has important implications for goal motivation. When induced to take a high-level view of a situation, people are more likely to evaluate their actions based on the desirability of the end state

(or goal) they hope to achieve rather than the time and effort required to achieve it (Liu, 2008; Rogers & Bazerman, 2008; Trope & Liberman, 2003). As a result, high-level thinking leads people to make choices that are more oriented towards goal achievement (Liberman & Trope, 1998; Liu, 2008; Trope & Liberman, 2003). We therefore predict that when temporal landmarks serve as interruptions, leading people to take a higher-level, big picture view of their lives, people's motivation to achieve their aspirations will increase.

Hypothesis and Study Overview

Integrating the past literature described above, we propose that temporal landmarks (a) separate people from their past imperfections and (b) shift people to think at a higher level about their lives and decisions. Consequently, we hypothesize that people will exhibit an increased tendency to pursue their aspirations following temporal landmarks.

Across three field studies, we test the hypothesis that temporal landmarks motivate aspirational behaviors, but that these effects weaken as people perceive themselves to be further from a temporal landmark. Based on past research on landmarks in autobiographical memory, we know that the beginning of a generic calendar cycle (e.g., the beginning of a week, month, or year), the beginning of a new period on an academic or work calendar (e.g., the first month of a semester, the first workday after a meaningful holiday) and the beginning of a new period in one's personal history (e.g., immediately following a birthday) serve as salient temporal landmarks (Robinson, 1986; Soster et al., 2010). We therefore predict that aspirational behaviors will increase

following these temporal landmarks.[2] The aspirational behaviors we examine primarily involve the initiation of behaviors that contribute to achieving a goal and tend to require repeated effort (e.g., dieting, exercising). Specifically, Study 1 uses daily Google searches for the term "diet" to examine how public interest in one particularly common aspirational activity changes over time. Study 2 tests whether *actual* engagement in an aspirational behavior (exercise) increases following temporal landmarks using university gym attendance records. Study 3 investigates the frequency with which people commit to a broad set of goals on the goal-setting website, "stickK" (www.stickk.com). Our findings are consistent with the hypothesis that we propose based on the theories described above. Although the current research primarily focuses on illustrating an important phenomenon and does not provide a direct test of the underlying mechanisms, these three field studies rule out a number of uninteresting alternative explanations for our findings, which we will discuss in the sections below.

Study 1: Google Searches for "Diet"

In Study 1, we measure public interest in the adoption of one aspirational behavior at different points in time. Specifically, we explore whether Internet searches for the term "diet" by the general population increase following temporal landmarks. Maintaining a healthy diet is considered one of the most effective methods for maintaining an optimal body weight (Shai et al., 2008), and about two-thirds of adult Americans are currently classified as overweight or obese (Centers for Disease Control

---

[2] Note that we examine the impact of a set of temporal landmarks that past research has shown demarcate the transition to a new mental accounting period. However, we do not address precisely what types of temporal landmarks produce fresh starts and what types of temporal landmarks fail to do so in the current research. This is a question worthy of future investigation.

and Prevention, 2013), making dieting an important goal for most Americans. Indeed,

dieting, losing weight, and eating more healthfully are among the most popular New

Year's resolutions listed on U.S. government's website (USA.gov, 2013). As described

above, we propose that temporal landmarks motivate the pursuit of aspirations by making

an individual feel segregated from and superior to her past, imperfect self and by

triggering her to take a big-picture view, which promotes a focus on goal attainment.

Therefore, we predict that people will search for the term "diet" more frequently

following temporal landmarks than on other days but that this increase will fade as the

temporal landmark recedes into the past.

Data

We obtained data from "Google Insights for Search"

(http://www.google.com/insights/search), a website where it is possible to download the

daily number of Google web searches that include a given search term dating back to

2004. Daily data on a given search term can only be extracted in intervals of three months

or less. We downloaded data on the daily number of Google searches in the United States

for the term "diet" over three-month intervals ranging from January 1, 2004 to June 30,

2012 (a time period including 3,104 days). Daily search volume data provided by Google

Insights for Search is both normalized relative to the total number of *daily* searches (for

any and all terms) on Google and further scaled based on search activity for the specific

query in question over the time period extracted (three months in this case). More

specifically, the day in a downloaded extraction period with the highest number of

searches (relative to total Google queries) is assigned a scaled value of 100, and other

days receive values that are scaled accordingly to fall between 0 and 100.[3] The relative

daily search volume ranges from 19 to 100 during the study period ($M = 64$, $SD = 18$).

See Appendix A for Google's description of these data.

Analysis Strategy

We examine whether people are more interested in dieting following temporal

landmarks using ordinary least squares (OLS) regression analyses. Our regression models

predict daily Google search volume for the term "diet" as a function of a series of

temporal landmark predictor variables described below. We estimate these regressions

with fixed effects for the 34 three-month intervals in our data to account for the fact that

search data is scaled within each interval and therefore cannot be compared directly over

time. We also cluster standard errors at the three-month interval level.[4]

Because public holidays and the start of a new week, a new month, and a new

year all represent partitions on the calendar, we expect that Internet searches for the term

"diet" will be highest immediately following these temporal landmarks. Notably,

individuals are naturally aware of the (continuously measured) day of the week (Monday-

Sunday), day of the month (1-31), and month of the year (1-12), which means they are

always aware of the time elapsed since the last temporal landmark corresponding to a

new week, month or year. However, calendars do not track the number of days that have

elapsed since the latest holiday. Thus, we do not expect people to be aware of how many

---

[3] Further, Google Insights for Search reports a search volume of "zero" when actual volume falls below a certain, undisclosed threshold. Zeros appear on seven days in our 3,104-day dataset. To ensure that these zero values did not spuriously magnify differences in search volume over time, we replaced each zero value with the lowest observed non-zero search frequency during the same extraction period. However, all reported results are robust to retaining zeros in our dataset.

[4] Our results do not change qualitatively or in terms of statistical significance if standard errors are not clustered.

days have elapsed since the last holiday they celebrated, but we do expect them to be

aware of how far they are from weekly, monthly, and yearly fresh start moments on the

calendar. In light of this, the predictor variables in our OLS regressions include measures

of a given day's distance from the beginning of the week, month, and year. However,

when evaluating the fresh start effect associated with public holidays, we simply test

whether searches for "diet" spike on the first workday after a holiday compared with

other, mundane days. Specifically, we include the following predictor variables in our

regression analyses to test for evidence of a fresh start effect:

- *Days since the start of the week.* We construct a continuous predictor variable

  indicating the days elapsed since the beginning of the current week (from 1 =

  Monday to 7 = Sunday).

- *Days since the start of the month.* We create a continuous predictor variable

  indicating the days elapsed since the beginning of the current month (min = 1,

  max = 31).

- *Months since the start of the year.* We include a continuous predictor variable

  indicating the number of months elapsed since the beginning of the current year

  (from 1 = January to 12 = December).[5]

- *First day after a Federal holiday.* We focus on the most widely celebrated U.S.

  holidays, or *Federal holidays,* which we define as one of the ten annual U.S.

  Federal holidays. We define the *first workday after a Federal holiday* as the first

---

[5] Because calendars count the time elapsed since the start of the year in months rather than days, we
specified our regressions accordingly, but notably our results are robust to instead including a measure of
the days elapsed since the start of the year.

day when people return to work after a *Federal holiday* and include a dummy

variable in our regressions to indicate whether or not a given day is the *first*

workday after a Federal holiday.

- *First workday x Fresh start score of Federal holiday.* If, as hypothesized,

  temporal landmarks elicit fresh start feelings and increase aspirational behavior,

  we would expect search volume for the term "diet" to be particularly high on days

  that feel more like a fresh start. For a separate research project, we identified a list

  of 26 holidays, 10 of which were the Federal holidays studied here. We asked 52

  participants on Amazon's Mechanical Turk to rate the extent to which each of

  these 26 holidays (or the day after it) felt like a fresh start on a seven-point scale

  (1 = not at all; 7=very much) (see Appendix B for these 26 holidays and the exact

  wording of our question). For the current study, we examine ratings of the 10

  Federal holidays of interest. For each of these 10 holidays, we averaged

  participants' ratings to form a composite fresh start score and standardized this

  score across the 10 holidays in our sample. We then created the variable *first*

  workday x fresh start score of Federal holiday by assigning the standardized

  rating of fresh start feelings associated with each Federal holiday to the first

  workday after a corresponding Federal holiday and assigning 0 to other days.

  Note that all reported results are robust to studying the set of 26 holidays rated

  instead of focusing only on the 10 Federal holidays.

Results

As predicted, we find that searches for the term "diet" are most frequent at the

start of each new calendar cycle: the beginning of the week, month, and year (see Model

1 in Table 1). First, searches for the term "diet" are more common at the beginning of the week and decrease as the week proceeds, as indicated by a significant, negative coefficient on *days since the start of the week.* Further, the significant, negative coefficients on *days since the start of the month* and *months since the start of the year* indicate that search volume for the term "diet" decreases over the course of each month as well as each year.

As we hypothesized, there is also an increase in search volume for the term "diet" following Federal holidays (see Model 1 in Table 1). Consistent with our prediction that temporal landmarks stimulate increases in aspirational behavior, there are more searches for "diet" following Federal holidays perceived as more like a fresh start. Specifically, a one standard deviation increase in a Federal holiday's fresh start rating is associated with a 6.78 point increase in daily search volume for the term "diet" (on a scale ranging from 0-100; $p < .001$, see Model 1 in Table 1).

Figure 1 illustrates that the magnitude of these effects is quite large when compared to the effect of the *New York Times* releasing a report on the successful clinical trial of an experimental diet pill in May, 2005 (Pollack 2005), a benchmark event that we expected to dramatically alter searches for the term "diet" (and which indeed increased "diet" search volume; $p < .001$). For example, the increase in daily search volume for the term "diet" associated with the start of the week (versus the end of the week) is about three times as large as the increase in search volume caused by this *New York Times* article.

Search volume for placebo terms. It is important to highlight that search volume for the term "diet" is already scaled by Google Insights for Search to adjust for the total number of daily Google queries, so the detected relationships between daily searches for "diet" and temporal landmarks cannot be attributed to changes in Internet search volume. However, to further exclude the possibility that our findings in Study 1 can be attributed to general patterns of Internet search over time, we compare searches for the term "diet" with searches for two popular search terms: "news" and "weather" (e.g., "news" was on Google's list of "hot searches" in the United States on July 23rd, 2012), which do not relate to aspirational behaviors. Furthermore, to empirically address two alternative explanations that may account for our findings in this paper (discussed in detail in our General Discussion), we identified and analyzed two additional placebo terms: "laundry" and "gardening." [6] We download daily search volume for these four terms during the same period when searches for the term "diet" are analyzed (from January 1, 2004 to June 30, 2012). When we re-run our models with the aforementioned placebo terms (news, weather, laundry, and gardening), we neither predict nor find that searches for these terms systematically increase following the temporal landmarks examined in Model 1 (see Models 2-5 in Table 1).[7]

---

[6] See the General Discussion section for details about the two alternative accounts as well as how we identified these placebo terms.

[7] Across our regressions with these four placebo terms, a few coefficient estimates are statistically significant in the predicted "fresh start" direction, whereas others show significant effects in the opposite direction. Consistent with our hypothesis, we did not observe reliable increases following temporal landmarks in searches for any of these placebo terms—only for the term "diet." The coefficient on days since the start of the week is a negative and significant predictor of daily searches for "news." A closer examination reveals that the negative coefficient on days since the start of the week, however, is driven by a dramatic drop in "news" search volume on weekends compared with weekdays, rather than by a gradual decline over the course of a week as is the case with searches for "diet" (and as the fresh start hypothesis predicts). In fact, people are significantly more likely to search for "news" on each day from Tuesday to

Discussion

The findings presented in Study 1 support our hypothesis that public interest in one important aspirational behavior – dieting – is higher following temporal landmarks. Specifically, we find that relative to baseline (Model 1 in Table 1), interest in dieting increases at the start of a new week (by 14.4%), a new month (by 3.7%), and a new year (by 82.1%), and following Federal holidays (by 10.2%). The effects cannot be attributed to general patterns of Internet traffic since the data we analyze is already scaled to account for overall search traffic on a given day and the search volume for other popular terms (news, gardening, laundry, and weather) does not exhibit the same systematic patterns.

Study 1 examines people's tendency to search for information about one particularly common aspirational behavior. However, we predict that the fresh start effect alters not only searches for information, but also actual decisions, as motivations and intentions are the first steps toward initiating actions and are predictive of behaviors (Ajzen, 1991; Gollwitzer, 1999). Our next study examines this prediction.

Study 2: Undergraduate Gym Attendance

By creating a discontinuity in our time perceptions and experiences, temporal landmarks can both psychologically separate individuals from their past imperfections and promote high-level thinking. Such processes are predicted to spur people to pursue aspirational behaviors following temporal landmarks. This is a hypothesis that we test in Study 2 by examining the frequency of engagement in one important aspirational

---

Friday relative to Monday, whereas people are more interested in dieting on Mondays than on all other days of the week (all p's < 00001; see Models A1 and A2 in Appendix C).

behavior – exercise. Increasing the frequency of exercise is one of the three most popular New Year's resolutions (Norcross et al., 2002; Schwarz, 1997). Like dieting, regular physical activity helps with weight loss and weight maintenance (Catenacci & Wyatt, 2007). However, only about 50% of American adults exercise as often as recommended by government guidelines (Centers for Disease Control and Prevention 2007). Thus, for many, exercise is an important but difficult-to-engage-in aspirational behavior.

In addition to examining actual engagement in an aspirational behavior (exercise), Study 2 also explores an additional, important predictor variable that was not available in Study 1. Specifically, in Study 2, we are able to investigate the impact on exercise of both calendar markers (e.g., holidays, the start of a new week, month or year) and one type of *personal* temporal landmark: birthdays.

Data

We obtained historical, daily gym attendance data for every undergraduate member ($N_{members}$ = 11,912) of a fitness center affiliated with a large university in the northeastern United States from September 1, 2010 through December 9, 2011 (N_{days} = 442).[8] Attendance was recorded automatically when students presented a magnetic student identification card to enter this facility. We also obtained information about the birthdates of a subset of these undergraduate members ($N_{members\_with\_birthday\_data}$ = 2,076). The number of students visiting the gym per day ranged from 31 to 2,270 during the study period (M = 883, SD = 470).

Analysis Strategy

---

[8] During this period, the gym was closed on 19 days. No observations about these days were therefore included in the raw dataset that the fitness center shared with us, and we thus exclude them from our analysis.

We conduct two types of OLS regressions to analyze our gym attendance data. The first aggregates attendance records across all undergraduate gym members on a daily basis. The outcome variable in this regression specification is the total number of gym visits on a given day divided by the number of hours the gym was open on that day (or the average gym visits per hour), which ranged from 5 to 142 ($M = 54$, $SD = 27$) in our sample. Our second analysis examines the likelihood that a given gym member visits the gym on each day in our dataset using an OLS regression model including fixed effects for each gym member and clustering standard errors at the date level.[9] The inclusion of gym member fixed effects controls for the effects of individual differences in time-invariant characteristics (e.g., gender, race, birth month) on gym attendance. To conduct this second analysis, we create a data set that contains one observation for each gym member on each day ($N_{person\text{-}days} = 5,265,104$). The dependent variable in this analysis equals one if a given gym member visited the gym on a given day and equals zero otherwise. In both of our regression specifications, we include predictor variables capturing the relationship between a given calendar day and temporal landmarks, as described below.

We predict that students will be more likely to visit the gym immediately following calendar landmarks and that their attendance will decline as these time markers become less salient. As in Study 1, we include *days since the start of the week, days since the start of the month,* and *months since the start of the year* as predictor variables in our

---

[9] We use an ordinary least squares regression model (rather than a more computationally intensive logistic regression model) because we include a large number of fixed effects and because logistic regression models typically produce inconsistent estimates when fixed effects are included unless data characteristics meet a stringent set of assumptions (Wooldridge, 2010). However, we obtain qualitatively similar results when we re-run our analyses using logistic regression models, though the significance of some predictors changes.

regressions. However, unlike in Study 1, we do not expect Federal holidays to be particularly salient calendar markers in the Study 2 student population because the university whose fitness center provided data for our study only closes for a subset of public holidays and has its own break schedule during the academic cycle. Thus, we expect the set of holidays and breaks recognized by this university to be more relevant landmarks than Federal holidays for our study population. As explained in Study 1, people do not naturally track the number of days elapsed since a recent holiday, so we measure the effects of holidays by creating a dummy predictor variable to indicate whether or not a day is the first day after any of the breaks listed on the university's academic calendar.

In addition, we expect the start of a new academic semester and birthdays to be meaningful partitioning points in the lives of the students included in our gym dataset. We predict that gym attendance will be highest immediately following the outset of a new semester and following an individual's birthday and will decline as the new semester or year of life proceeds. We include the following predictor variables in our regression analyses to test these hypotheses:

- *Months since the start of the semester.* We include a continuous predictor variable indicating the months elapsed since the beginning of the current semester (e.g., 1 = September or January; 4 = December or April).

- *Months since last birthday.* We were able to obtain information about the birthdates of a subset of 2,076 gym members, which we matched with their gym attendance records. We define a *birth year* as a personalized year that starts on the

first day following an individual's birthday and ends on his or her next birthday. For each of the 2,076 students in our dataset with a known birthday, we include a continuous predictor variable in our regressions indicating the months elapsed since their last birthday. Specifically, we calculate the distance in days between each date in the study period and a given student's previous birthday. We convert this distance to units of months with each "month" taking on the actual length of the appropriate calendar month (e.g., 1 = the 31 days immediately following an individual's birthday; 12 = the 31 days immediately preceding an individual's birthday, including the birthday itself).

We control for a number of other variables that may affect a student's likelihood of attending the gym. Since college students are likely to be away from campus during school breaks, we create one dummy variable to indicate whether the university studied was in normal class session (fall and spring semesters, excluding school breaks) and another dummy variable to indicate whether the university was in summer session on a given date. Furthermore, since exam periods occur at the end of the semester and the calendar year, it is possible that *month of the year* and *month of the semester* affect gym attendance because students are busier than usual or more likely to have left school during exams. To alleviate this concern, we control for whether each date fell during the university's final exam period. To account for the fact that more students leave campus as the exam period progresses, we also include a variable in our regressions to indicate the number of days since the start of the final exam period, which is coded as zero for dates falling outside of the university's final exam period. All reported results are also robust to

excluding days falling during exam periods from our data analysis. For the analyses at the level of the individual gym member, we also control for the number of hours that the gym was in operation on a given calendar date.

Results

Models 6-8 in Table 2 present results from OLS regressions exploring the statistical relationship between temporal landmarks and (a) average gym visits per hour across all gym members (Model 6) and (b) daily gym attendance by individual members (Models 7 and 8).

First, as we observed with searches for the term "diet," we find that gym attendance increases at the start of each new week, month, and year. As Models 6 and 7 in Table 2 show, *days since the start of the week* takes on a significant, negative coefficient, indicating that people visit the gym less as each week proceeds.[10] Further, the significant, negative coefficients on *days since the start of the month* and *months since the start of the year* in Models 6 and 7 suggest that gym attendance decreases over the course of each month as well as each year.[11] In addition, as hypothesized, Models 6 and 7 show that students exercise more both at the start of a new semester (relative to the end of the semester)[12] and on the first day after a school break.

---

[10] In separate regressions where we replace days since the start of the week with six indicator variables – one for each day of the week from Tuesday to Sunday (with Monday omitted) – we find that hourly gym traffic is higher on Mondays than on all other days (all p's < 0.05 except the comparison with Tuesday; see Models A6-A8 in Appendix D).

[11] It is worth noting that students in our study do not pay to use the gym: all enrolled undergraduates are automatically granted memberships at the university's fitness facility. Therefore, the observed decrease in usage over the course of a given month or semester could not be attributed to gradually decreasing sensitivity to membership payments as described by Gourville and Soman (1998).

[12] Importantly, the finding that gym attendance decreases over the course of a semester, though consistent with our proposed fresh start effect, may be driven by the fact that students become busier as the semester proceeds. However, other temporal landmarks examined in this study could not be explained by this alternative account.

For the subset of 2,706 gym members whose birthdates were made available to us, we explore whether the likelihood that a student visits the gym is higher in the weeks and months immediately following a birthday than later in the year. In an initial regression analyzing daily gym attendance in this sub-population, we actually observe a positive correlation between the variable *months since birthday* and gym attendance ($\beta = 3.6*10^{-4}$, $p < 0.0001$) – the opposite of our prediction. However, when we examine this relationship more closely, we find that gym members react dramatically differently to their 21st birthdays than to other birthdays. Specifically, students turning 21 tend to decrease their gym attendance following this birthday. However, for students celebrating other birthdays, we observe the predicted, significant and negative correlation between *months since birthday* and gym attendance (see Model 8 in Table 2). This indicates that students exercise more frequently right after most birthdays. The 21st birthday exception may be related to the fact that this birthday corresponds to the date when students are first legally permitted to purchase alcoholic beverages or to the fact that it is associated with an increase in autonomy and social status, which may reduce students' urges to change themselves for the better. Of course, while it is interesting that the 21st birthday is qualitatively different from other birthdays, it is important to highlight that potential explanations are entirely speculative.

To confirm that the 21st birthday differs significantly from other birthdays with respect to the predicted fresh start effect, we ran a regression including observations of all students with available birthdate data to predict whether each student visited the gym on each date in our dataset. We added a dummy variable (*age 21*) to indicate whether an

observation corresponded to a day in the year following a gym member's 21st birthday and interacted this dummy variable with all other predictor variables in our model (including control variables and person fixed effects). The interaction between *months since birthday* and *age 21* is significant and positive in this model ($\beta = 7.9*10^{-4}$, $p < 0.05$), which means that the coefficient on *months since birthday* for observations associated with all birthdays other than the 21st is significantly larger than the coefficient for observations associated with students' 21st birthdays.[13] Because we are interested in the effect of birthdays on gym attendance at a typical age, we report the results from analyses of all other birthdays in Table 2 (see Model 8). In a regression where we replace *months since birthday* with 11 dummy variables to indicate each month in a person's birth year (with the half-birthday month marker as the omitted reference month), we find that people are more likely to exercise during the first month after a birthday than during the half-birthday month ($\beta = 2.9*10^{-3}$, $p < 0.05$), and they are also less likely to exercise during the final month preceding a birthday ($\beta = -3.4*10^{-3}$, $p < 0.05$). We conclude that birthday temporal landmarks typically motivate exercise, and motivation declines over the course of the year, reaching its lowest level in the final month preceding a birthday.

Figure 2 illustrates the magnitude of these effects. Specifically, these effects are compared to the impact of extending the gym's hours of operation by one hour (which itself is a significant, positive predictor of attendance; $p < 0.001$). We observe that the effects of temporal landmarks on gym attendance are quite large in comparison with the effect associated with extended hours. For example, the increase in an individual's probability of going to the gym in the month immediately following a birthday (versus

the month immediately preceding a birthday) is equivalent to the effect associated with keeping the gym open for two extra hours.

Discussion

Study 2 shows that people are more likely to exercise following temporal landmarks: the probability of visiting the gym increases at the beginning of a new week (by 33.4%), month (by 14.4%), year (by 11.6%), and semester (by 47.1%), as well as following school breaks (by 24.3%), relative to baseline (Model 7 in Table 2). In addition to replicating the findings of Study 1 with a consequential behavioral outcome, Study 2 also demonstrates that personally-relevant temporal landmarks – namely, birthdays – are, like calendar landmarks, associated with subsequent upticks in aspirational behavior. In this case, the probability of visiting the gym is increased by 7.5% following birthdays besides the 21st (Model 8 in Table 2).

One alternative explanation for some of our findings in Studies 1 and 2 is that people consume a larger amount of food on certain temporal landmarks, such as holidays and weekends. As a result, people might try to reduce their caloric intake or exercise more intensively following these "binges" in an attempt to lose weight gained leading up to temporal landmarks. This alternative account suggests that the tendency to start healthier routines following temporal landmarks is simply a *physiological* response to the health effects of overindulgence. In light of the concern that some Federal holidays are excuses for gluttony, we conducted robustness checks by removing Independence Day, Labor Day, Thanksgiving Day, and Christmas from the list of public holidays and school breaks included in our regression analyses. We found that daily Google searches for the term "diet," average gym visits per hour, as well as the probability of visiting the gym are

still significantly higher on the first workday after a Federal holiday or a school break than on typical days. In spite of this alternative explanation's inability to account for all of our empirical findings in Studies 1 and 2, to more carefully address the possibility that the fresh start effect is exclusively the product of overeating on weekends and holidays, we conduct an additional study.

<div align="center">**Study 3: Commitment Contracts**</div>

The objective of Study 3 is to demonstrate that following temporal landmarks, people take steps to tackle a broad set of goals that they aspire to achieve, and increases in the intensity of goal pursuit cannot be explained by the physiological alternative explanation articulated above. We expect temporal landmarks to propel the pursuit of a broad set of goals because temporal landmarks, by demarcating new mental accounting periods, can both psychologically distance the current self from past imperfections and direct an individual to focus on high-level, goal-relevant ambitions.

Data

We obtained data from stickK (www.stickK.com), a website that helps customers achieve their personal goals. Specifically, stickK offers users an opportunity to set personal goals and specify consequences that will ensue if they fail to achieve those goals. It is well-documented that goal-setting establishes reference points (e.g., Heath, Larrick, & Wu, 1999; Sackett, Wu, White, & Markel, 2012) and is instrumental to goal achievement (Locke & Latham, 1990). To create what stickK terms a "commitment contract," users first specify their goal and select a date by which they contractually agree to accomplish it. Next, users choose an amount of money to forego if they fail to achieve their goal. When users put a positive amount of money on the line, they also select a

recipient of these stakes (e.g., a friend, a charity), should they fail to achieve their goal. Finally, users have the option to (a) designate a third party to monitor and verify their achievements and (b) designate other stickK users as their supporters. When creating a commitment contract, users can choose one of stickK's five standard goals (*exercise regularly, lose weight, maintain weight, quit smoking, or run a race*) or specify a *custom goal*, which they are asked to classify into one or more of the following categories: career, diet and healthy eating, education and knowledge, exercise, family and relationships, green initiatives, health and lifestyle, home improvement, money and finance, personal relationships, quit smoking, religion, hobbies and recreation, and weight loss.

The data that stickK provided for this study contains 66,062 records of commitment contracts that 43,012 unique users created between October 1, 2010 and February 13, 2013 (N_{days} = 886). Table 3 lists summary statistics for the number of contracts created per day in each goal category.

Analysis Strategy

We conduct two types of analyses with data from stickK. The first aggregates commitment contracts across all users on a daily basis ($N_{contracts}$ = 66,062, N_{users} = 43,012, N_{days} = 866) and relies on OLS regression models to predict the total number of contracts created each day. The second method allows us to examine the motivating effects of birthdays by examining the likelihood that a given user creates a goal on each day in our dataset using an OLS regression model including fixed effects for each of 42,913 users whose birthdates were made available to us. We cluster standard errors at the date level. As in Study 2, we create a data set that contains one observation per user per day ($N_{person-}$

$_{days}$ = 37,162,658). We set the dependent variable in this person-day analysis equal to one if a given stickK user created a commitment contract on a given day and zero otherwise.

As in Studies 1 and 2, we create a set of predictor variables indicating a given calendar day's proximity to the beginning of the week (*days since the start of the week*), the beginning of the month (*days since the start of the month*), and the beginning of the year (*months since the start of the year*). Using the same methods described in Study 1, we construct (a) the dummy variable, *first workday after a Federal holiday,* to indicate whether a given day is the first workday after a Federal holiday, and (b) the variable *first workday x fresh start score of Federal holiday*, to indicate the extent to which each Federal holiday was rated as a fresh start. We again expect that birthdays represent important personal temporal landmarks and therefore promote a focus on aspirations. For the subset of 42,913 stickK users whose birthdates were available to us, we create the additional predictor variable *months since last birthday* using the method described in Study 2 ($N_{contracts}$ = 65,845).

We control for stickK's considerable growth in users and contracts during our study period. For each calendar day in our dataset, we create a control variable, *days since launch*, to indicate the number of days elapsed since the start of our dataset. We also include its quadratic term to control for the potentially non-linear trend in the growth of the stickK customer population.[13]

Results

---

[13] Our independent variables of interests remained qualitatively the same in terms of magnitude and statistical significance if we excluded the quadratic term from our regression models.

All types of commitment contracts. Consistent with our hypothesis, we find that goal contracts are created more frequently at the beginning of the week than at the end of the week, as indicated by a significant, negative coefficient on *days since the start of the week* in our regression models (see Models 9 and 10 in Table 4).[14] Also, the significant, negative coefficients on *days since the start of the month* and *months since the start of the year* in Models 9 and 10 in Table 4 indicate that people create more new goals at the beginning of the month and year as compared with the end of the month and year. Further, as we hypothesized, the total number of commitment contracts increases immediately following Federal holidays, and the magnitude of this increase is larger after holidays rated as more likely to elicit fresh start feelings (see Models 9 and 10 in Table 4).

We next turn to an exploration of whether the likelihood that a user creates a contract is higher in the weeks and months immediately following his or her birthday compared with later in the year. For the 42,913 users in our dataset with a known birthdate, the variable *months since birthday* is a negative and marginally significant predictor of the likelihood that a user will create a goal contract on a given day (see Model 10 in Table 4). This suggests that people are more motivated to pursue goals following a birthday than preceding one.[15] In a regression where we replace *months since birthday* with 11 dummy variables to indicate each month in a person's birth year (with

---

[14] In a separate regression where we replace *days since the start of the week* with six indicator variables – one for each day of the week from Tuesday to Sunday (with Monday omitted) – we find that the number of commitment contracts created is significantly higher on Mondays than on any other day of the week (all *p*'s <0.05 except the comparison with Tuesday; see Models A9 and A10 in Appendix E).

[15] Note that we do not expect the 21st birthday to differ from other birthdays (and do not find that it differs) when it comes to general goal setting. The legal option to purchase alcohol may alter one's immediate inclination to exercise (Study 2) but it should not affect one's general inclination to set goals (Study 3).

the half-birthday month marker as the omitted reference month), we find that people are significantly more likely to create a commitment contract during the first month after a birthday than during the month of their half-birthday ($\beta = 7.0*10^{-5}$, $p < 0.05$), and they also show an (insignificant) trend of creating fewer commitment contracts in the last month before their birthday relative to their half-birthday month ($\beta = -3.4*10^{-5}$, $p > 0.10$).

Figure 3 illustrates that the magnitude of these effects is quite large in comparison with the impact of *ABC News* releasing a feature article about stickK in March, 2012 (Farnham 2012), a benchmark event that we would expect to dramatically increase attention to stickK (indeed, this article significantly increased the number of contracts created on the day of its release; $p < 0.05$). For example, the increase in an individual's probability of creating a contract right after a Federal holiday (relative to other more mundane days) is four times as large as the effect of the release of this *ABC News* article.

Commitment contracts for custom goals. It is important to address the possibility that some of our findings in Studies 1 and 2 could be driven by over-indulgence associated with certain types of temporal landmarks (e.g. holidays, weekends, birthdays), which might lead to subsequent compensatory exercising and dieting. To address this possibility, we investigate the patterns described above for *custom goals* that are not health-related. As described above, when creating a custom goal, stickK provides a list of goal sub-categories and requires users to check all sub-categories that apply. The list of sub-categories encompasses a broad set of domains, including many that are not directly related to health (specifically, these include: career, education and knowledge, money and finance, personal relationships, green initiatives, home improvement, religion,

family and relationships, as well as hobbies and recreation). Examples of health-irrelevant custom goals that are featured on www.stickK.com include "being on time", "spending more time with family", "helping others", "learning something new", and "reducing debt" (http://www.stickk.com, accessed July 28, 2013). To ensure that the fresh start effect is not simply the result of compensatory cutbacks following overindulgence, we focus on *custom goals* for which stickK users did not select any health-related sub-categories ($N_{contracts}$ = 15,213, N_{days} = 866, N_{users} = 10,074). Using the same OLS regression model specifications described above in the Analysis Strategy section of Study 3, we predict the total number of contracts created each day for health-irrelevant custom goals.

As predicted, health-irrelevant custom goal contracts (see Table 4, Models 11 and 12) are created more frequently at the beginning of the week, month, and year, following Federal holidays, and particularly after holidays rated as more like a fresh start, compared with other days. Although there is a trend whereby more health-irrelevant custom goals are initiated following a birthday, this trend is not significant (see Table 4, Model 12). Models 13-15 in Table 5 report regression results for the three most popular health-irrelevant custom goals (career, education and knowledge, and money and finance), which all show these same trends.

Robustness across goal types. We find the same basic patterns of results when we separately analyze health-relevant custom goals as well as the five types of standard goal contracts offered by stickK: *exercise regularly*, *lose weight*, *maintain weight, quit*

smoking, and *run a race*. See Models 16-21 in Table 5 for regression results broken down by goal type.

Discussion

Consistent with our hypothesis, Study 3 shows that relative to baseline, people are more likely to commit to their goals at the beginning of a new week (by 62.9%), month (by 23.6%), or year (by 145.3%), and following Federal holidays (by 55.1%), as well as following their birthdays (by 2.6%) (Model 10 in Table 4). Further, Study 3 provides evidence that the fresh start effect pertains to a broad set of health-irrelevant goals (e.g., career, education and knowledge, and personal relationships). This suggests that the increase in aspirational behaviors following temporal landmarks that we document throughout this paper cannot be parsimoniously explained by the physiological need to offset overindulgence.

General Discussion

Across three field studies, we find evidence of a fresh start effect whereby people exhibit a higher likelihood of engaging in aspirational behaviors following temporal landmarks such as the initiation of new calendar cycles (e.g., the start of a new week, month, year, or academic semester), holidays, and birthdays. We analyze a broad set of aspirational activities: web searches for the term "diet", gym attendance, and the creation of commitment contracts to support a wide range of different goals. The effects we document are large in magnitude, suggesting that the fresh start effect has meaningful implications for individual and societal welfare.

The fresh start effect documented in this paper is consistent with two psychological processes we proposed to parsimoniously explain it. First, new mental

accounting periods as demarcated by temporal landmarks psychologically distance the current self from past imperfections, propelling people to behave in line with their new, positive self-image. Second, temporal landmarks interrupt attention to day-to-day minutiae, causing people to take a big-picture view of their lives and thus focus more on achieving their goals. This paper relies on field data to demonstrate the existence of the fresh start effect, but it does not offer a direct test of the underlying mechanisms responsible for this effect. Thus, future research documenting the psychological processes that underlie the fresh start effect would be extremely valuable. In the next section, we discuss and provide evidence that helps rule out a number of uninteresting alternative explanations for our findings.

Alternative Explanations

One concern with our findings is that people tend to engage in activities prior to (or during) temporal landmarks that harm goal pursuit, and our findings might simply reflect a rational attempt to offset these bad behaviors *after* temporal landmarks. For example, the fresh start effect could simply be attributed to the desire to counteract excessive caloric intake associated with weekends and holidays. We can rule out this alternative explanation in a number of ways. First, in Study 3, we rule out this alternative explanation by showing that following temporal landmarks, commitment contracts for health-irrelevant goals increase. Second, when we remove holidays that are particular excuses for gluttony (Independence Day, Labor Day, Thanksgiving, and Christmas), we still find a significant uptick in aspirational behaviors immediately following holidays and school breaks. Third, this compensatory alternative explanation cannot account for our consistent finding that aspirational behaviors are more intense at the start of the

month than at the close of a month since neither the start nor the end of a new month is associated with increased indulgence. Finally, this alternative explanation suggests that engagement in aspirational activities would be significantly lower right before temporal landmarks than on other days. We can directly test whether this is the case by exploring whether people are indeed significantly less likely to engage in aspirational behaviors immediately before temporal landmarks than on other days across our three field data sets.

Although we hypothesize that temporal landmarks elevate the frequency of aspirational behaviors and that these effects weaken as people perceive temporal landmarks to be further away, our hypothesis does *not* predict that engagement in aspirational behaviors will be significantly lower in the short period immediately preceding (or during) a temporal landmark than on any other, typical day. Therefore, we created indicator variables for weekends, the last seven days of each month, the last seven days of each year, the seven days preceding the first workday after each Federal holiday (Studies 1 and 3), the seven days preceding the first school day after each school break (Study 2), the seven days preceding each semester's start (Study 2), and the seven days immediately before and including a person's birthday (Studies 2 and 3). We then added these additional predictor variables to our primary regression models (Models 1, 6, 7, 8, 9 and 10). If our findings were simply attributable to reduced engagement in aspirational behaviors prior to temporal landmarks, we would expect the coefficients on these new predictor variables to be significant and negative. In fact, among 29 new predictor variables across six regression models, only three predictor variables have a significant,

negative coefficient at the 5% level, which is not significantly more than the number that would be expected by chance. In addition, the inclusion of these predictor variables does not qualitatively change the coefficients on our primary predictor variables, which remain essentially the same in terms of magnitude and statistical significance. Therefore, it is unlikely that our findings are solely driven by people's reduced engagement in aspirational behaviors prior to temporal landmarks.

Another alternative explanation for our findings is that people do not have enough time and energy to tackle their goals before temporal landmarks and thus put off aspirational behaviors until after temporal landmarks have passed. Such an alternative account suggests that the period before a temporal landmark is not a good time to initiate goal pursuit and thus should be associated with a significant dip in the frequency of aspirational behaviors, but the analyses described above show that this is not the case. Further, while it is likely that the arrivals of some new mental accounting periods (e.g., following a wedding or a job change) are accompanied by more free time to tackle goals than the windows preceding them, people do *not* typically have more free time to pursue aspirational activities following most of the types of temporal landmarks studied in this paper (e.g., the beginning of a new week, the beginning of a new month, the first workday after a holiday, or during the first few months following a birthday) than before these temporal landmarks (e.g., on the weekend, at the end of the month, before or during a holiday, or in the few months preceding a birthday). To further address this alternative explanation, however, we recruited 53 participants online from Amazon's Mechanical Turk to participate in a survey about daily activities. They were first asked to list three

activities that they had the tendency to put off doing until a future date when they thought they would have more time and energy. Next, participants were asked to select the subset of activities from their list that were *not* aspirational (see Appendix F for the exact questions). A research assistant removed activities that fit our definition of "aspirational" and then identified the most frequently listed activity that participants tended to put off doing and that was not aspirational in nature: "laundry." Following the procedures described in Study 1, we downloaded daily Google search volume for this word from January 1, 2004 to June 30, 2012. We neither predict nor find that searches for "laundry" systematically increase following the temporal landmarks examined in Model 1 (see Model 4 in Table 1), suggesting that temporal landmarks do not simply increase the interest in all types of activities that require planning, time, and energy.

There are several other potential explanations for the documented fresh start effect besides the psychological processes we propose that can be ruled out. First, it could be argued that people generally embrace all types of new activities at the beginning of new cycles. Study 3 helps address this alternative account by showing that the fresh start effect is not confined to the adoption of new habits. For example, temporal landmarks are followed by an increase in the number of commitment contracts created for smoking cessation, an aspirational behavior that disrupts an existing habit (see Table 5). To further address this alternative account, we recruited another 49 participants online from Amazon's Mechanical Turk to list three "new" activities that they had never engaged in before but would consider pursuing in the future. As in the survey we described above, we again asked participants to indicate the subset of activities on their list that were not

aspirational (see Appendix F for the exact questions) and asked a research assistant to remove activities that fit our definition of "aspirational". "Gardening" was the most frequently listed "new" activity that was not aspirational in nature. We did not find that Google searches for "gardening" systematically increase following the temporal landmarks examined in Study 1 (see Model 5 in Table 1), suggesting that temporal landmarks do not induce increased engagement in all types of new activities.

It is also important to note that some temporal landmarks, particularly personally-meaningful life events (e.g., a wedding, a job change) tend to alter one's surroundings and daily routines, which in turn trigger certain habitual actions. Past research has shown that altering one's surroundings and routines can lead to behavior change (Wood, Tam, & Witt, 2005). For example, a move to a new residence may promote a healthy lifestyle because recurring stimuli that cue old, unhealthy habits no longer exist (e.g., a favorite bakery is now far away). Alternatively, a move to a new residence may promote an unhealthy lifestyle because a favorite salad shop is no longer nearby and instead an ice cream parlor is just down the street. There are several reasons why we believe we can rule out this explanation for our findings. First, while many temporal landmarks do disrupt routines, many of those we study (e.g., the start of a new week/month, the celebration of a birthday) do not typically alter routines significantly. In fact, weekly and monthly cycles may actually reinforce routines. Second, this past research on habit disruption does not clearly predict whether contextual shifts that may be induced by certain types of temporal landmarks will lead to increases in aspirational or harmful behaviors. In fact, there is evidence that routine changes can disrupt beneficial habits

such as reading the newspaper (Wood et al., 2005). Thus, past research on routines and habit formation does not seem likely to explain the fresh start effect detected in this paper.

It could be argued that some temporal landmarks associated with relaxation, such as weekends and holidays, might replenish self-regulatory resources, restoring the self-control that people need to tackle aspirational behaviors (Baumeister et al., 1998). Though repletion could contribute to the elevated motivation to pursue goals that we detect following weekends and holidays and strengthen the impact of the psychological processes highlighted in the Conceptual Framework section, this account cannot explain why people choose to engage in aspirational activities at a higher rate following the start of a new month or immediately following a birthday. Also, more nuanced analyses of our field data suggest that the observed fresh start effect is unlikely to be solely driven by changes in self-regulatory resources. Specifically, this alternative account predicts that the frequency of aspirational behaviors should be higher on Saturday and Sunday than Friday because having a day off from work or school is relaxing. However, eight regression models where we replace *days since the start of the week* with six indicator variables – one for each day of the week from Tuesday to Sunday (with Monday omitted) – provide no consistent evidence that Friday is associated with lower engagement in aspirational behaviors than either Saturday or Sunday (see Models A1, A6-8, and A9-12 in Appendices C, D, and E, respectively). In concurrent research exploring the mechanism underlying the fresh start effect through laboratory experiments, Dai, Milkman, and Riis (2015) show that people are more motivated to pursue aspirational

behaviors following more psychologically meaningful temporal landmarks (e.g., a meaningful birthday or job change) than objectively commensurate but less psychologically meaningful temporal landmark (e.g., a typical birthday or job change). These findings help rule out relaxation as the sole explanation for the fresh start effect because psychologically meaningful temporal landmarks would not be expected to provide greater opportunities for relaxation than objectively identical but less meaningful landmarks.

Implications

The fresh start effect has significant practical implications for individual decision makers, managers, and policy makers. First, individuals can not only take advantage of their fresh start feelings at naturally-arising temporal landmarks to follow through on good intentions, but they may also be able construct fresh starts themselves by strategically "creating" turning points in their personal histories, such as moving to a new residence to start over (a previously-named phenomenon called "relocation therapy"; Kaufman, 2013). Second, our findings suggest new ways in which people may be effectively "nudged" (Thaler & Sunstein, 2008) to begin pursuing their aspirations. For example, messages designed to promote aspirational behaviors may be most impactful at fresh start moments (e.g., the beginning of a new month, right after holidays) when message recipients will be more interested in striving to achieve their long-term goals, as shown in this paper. Further, marketers of products designed to help people attain desirable objectives (e.g., retirement counseling services, gym memberships, online education programs) may best appeal to consumers' desires for self-improvement by advertising at fresh start moments.

Another implication of this research is that framing certain days as opportunities for a fresh start (e.g., birthdays, the start of a new week/month/year, etc.) may help people make choices that maximize their odds of achieving their aspirations. For example, employers could potentially reframe transition points in the workplace (e.g., a desk move, or a return from vacation) to increase the adoption of aspirational activities (e.g., attending training workshops or onsite biometric screenings).

An important question related to the practical implications of fresh start effects is how long fresh-start feelings persist following the incidence of a temporal landmark. Plots (see Appendix G) suggest that the elevated motivation we document in this paper spikes on the first workday after a Federal holiday and declines rapidly thereafter, whereas motivation wears off much more gradually over the course of each week, month, year, and semester. However, it is worth noting that even fleeting fresh start feelings following temporal landmarks can potentially be valuable for at least two reasons. First, the abundance of fresh-start opportunities throughout the year offer repeated chances for people to attempt positive self-change, so even if they initially fail, they may subsequently succeed (Polivy & Herman, 2002). Second, transient increases in motivation may be sufficient to help people fulfill important one-shot goals such as receiving a medical test or signing up for a 401(k) account with monthly payroll deductions. In this paper, we primarily study aspirational behaviors where the end goal requires engaging in a series of goal-directed actions (e.g., dieting, exercising, committing to a personal goal). It would be valuable for future research to examine the

extent to which temporal landmarks can spur aspirational behaviors that only require a single action (e.g., getting a vaccine, donating to a charity).

Limitations and Future Directions

The empirical evidence presented in this paper primarily focuses on temporal landmarks associated with socially-constructed timetables (including the yearly calendar, work calendar, and academic calendar). Birthdays are the one exception and example of personally-relevant temporal landmarks studied here. Further, we focus on the Gregorian calendar given its relevance to the settings studied. Future research exploring and comparing a broader set of temporal landmarks, including temporal landmarks on different calendars (e.g., the Chinese New Year, the Jewish New Year) as well as additional personal landmarks (e.g., religious conversions, relocations, job changes, etc.) would be valuable. We expect that the fresh start effect likely extends to all temporal landmarks, not only those examined in this paper, though certain types of temporal landmarks may produce stronger effects than others (Dai et al., 2015).

In addition, the temporal landmarks highlighted here are all associated with either neutral or positive experiences. Temporal landmarks of negative valence (e.g., a divorce, the death of a family member) may not immediately increase motivation to pursue aspirations if people need to first cope with stressful experiences (Cohen & Hoberman, 1983). It would be valuable for future research to explore whether the fresh start effect extends to temporal landmarks stained by negative emotions such as grief and anger.

Our findings raise a number of other questions worthy of exploration. One such question is how the anticipation of a temporal landmark affects behavior. Some recent work suggests that people might feel less compelled to begin pursuing their goals when

upcoming landmark events are highlighted because the future self (who will benefit from goal pursuit) feels more disconnected from the current self (Bartels & Rips, 2010; Bartels & Urminsky, 2011; Tu & Soman, 2014). On the other hand, Peetz and Wilson (2013) contend that when an intervening landmark event and a future desirable state are both made salient, the discrepancy between the current self and the future, desired self is highlighted, which motivates beneficial behaviors. Another possibility is that people may use upcoming landmarks as self-imposed deadlines and attempt to bring ongoing goals to closure by these deadlines (e.g., finish reading a book, complete an assignment). Our research suggests two other possible effects of anticipating an upcoming temporal landmark. First, anticipated temporal landmarks might liberate people to make goal-incongruent choices if they anticipate wiping the slate clean after an upcoming temporal landmark (Zhang, Fishbach, & Dhar, 2007). Second, if a decision maker foresees that a better opportunity to pursue her aspirations will arise following an impending landmark (e.g., after her next birthday), she may strategically delay launching her plans until after the landmark. Future research exploring these possibilities would be valuable.

Further, future research could explore if and how social influence reinforces the fresh start effect. For example, a spike in goal pursuit on January 1 may partly reflect a social bandwagon effect. Though other fresh start moments highlighted in the current research (e.g., the beginning of the week or month) attract less attention, the fresh start effects we observe across three studies could be magnified in part by a social contagion process whereby others' engagement in aspirational activities stimulates increases in our own goal motivation. Exploring this hypothesis in future research would be valuable.

References

Ahn, H. K., Liu, M. W., & Soman, D. (2009). Memory markers: How consumers recall the duration of experiences. *Journal of Consumer Psychology, 19*(3), 508-516. doi: 10.1016/j.jcps.2009.05.002

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179-211. doi: 10.1016/0749-5978(91)90020-T

Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*(6), 1423-1440. doi: 10.1037//0022-3514.83.6.1423

"Aspiration". Merriam-webster.com. Retrieved July 29, 2013 from http://www.merriam-webster.com/dictionary/aspiration.

Ayers, J. W., Althouse, B. M., Johnson, M., & Cohen, J. E. (2014). Circaseptan (weekly) rhythms in smoking cessation considerations. *JAMA Internal Medicine, 174*(1), 146-148. doi: 10.1001/jamainternmed.2013.11933

Bartels, D. M., & Rips, L. J. (2010). Psychological connectedness and intertemporal choice. *Journal of experimental psychology. General, 139*(1), 49-69. doi: 10.1037/a0018062

Bartels, D. M., & Urminsky, O. (2011). On intertemporal selfishness: How the perceived instability of identity underlies impatient consumption. *Journal of Consumer Research, 38*(1), 182-198. doi: 10.1086/658339

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 74*(5), 1252-1265. doi: 10.1037/0022-3514.74.5.1252

Bazerman, M. H., Tenbrunsel, A. E., & Wade-Benzoni, K. (1998). Negotiating with yourself and losing: Making decisions with competing internal preferences. *Academy of Management Review, 23*(2), 225-241.

Bhargave, R., & Miron-shatz, T. (2012). *Forty and fabulous? Milestone agers assess life satisfaction more on income than positive*. Working Paper. The University of Texas at San Antonio, San Antonio, TX.

Botti, S., Broniarczyk, S., Häubl, G., Hill, R., Huang, Y., Kahn, B., . . . Wansink, B. (2008). Choice under restrictions. *Marketing Letters, 19*, 183-199.

Catenacci, V. A., & Wyatt, H. R. (2007). The role of physical activity in producing and maintaining weight loss. *Nature Clinical Practice Endocrinology & Metabolism, 3*(7), 518-529. doi: 10.1038/ncpendmet0554

Centers for Disease Control and Prevention. (2007). U.S. Physical activity statistics. Retrieved May 21, 2013, from http://apps.nccd.cdc.gov/PASurveillance/DemoComparev.asp.

Centers for Disease Control and Prevention. (2013). Faststats: Obesity and overweight. Retrieved July 29, 2013, from http://www.cdc.gov/nchs/fastats/overwt.htm

Cohen, S., & Hoberman, H. M. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology, 13*(2), 99-125. doi: 10.1111/j.1559-1816.1983.tb02325.x

Cross, A., Peretz, P., Munoz-LaBoy, M., Lapp, I., Shelley, D., & Rosenfield, A. (2006). Monday - a social framework to enhance health promotion efforts and facilitate sustainable health behavior change. Retrieved November 15, 2012, from www.dkv.columbia.edu/demo/healthy_monday/pdfs/white-paper.pdf

Dai, H., Milkman, K. L., & Riis, J. (2015). *Put your imperfections behind you: Why and how meaningful temporal landmarks motivate aspirational behavior*. Working Paper. Retrieved from http://ssrn.com/abstract=2420476

Epstein, S. (1973). Self-concept revisited or a theory of a theory. *American Psychologist, 28*(5), 404-416. doi: 10.1037/H0034679

Farnham, A. (2012). Got goals but lack discipline? Now you can buy it online. ABC News. Retrieved March 14, 2013, from http://abcnews.go.com/Business/stickk-websites-pay-cash-meet-goals/story?id=15904736#.UbISLJwV85J.

Fry, J., & Neff, R. (2010). *Healthy monday: Two literature reviews*. Working Paper. The Johns Hopkins Center for a Livable Futurem, Baltimore, MD.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*(7), 493-503. doi: 10.1037/0003-066X.54.7.493

Gourville, J. T. (1998). Pennies-a-day: The effect of temporal reframing on transaction evaluation. *Journal of Consumer Research, 24*(4), 395-408. doi: 10.1086/209517

Heath, C., & Soll, J. B. (1996). Mental budgeting and consumer decisions. *Journal of Consumer Research, 23*(1), 40-52. doi: 10.1086/209465

Kaufman, J. (2013). Relocation therapy. *The New York Times*. Retrieved January 4, 2013, from http://www.nytimes.com/2013/01/06/realestate/relocation-therapy.html?pagewanted=all&_r=0.

Khan, U., & Dhar, R. (2006). Licensing effect in consumer choice. *Journal of Marketing Research, 43*(2), 259-266. doi: 10.1509/jmkr.43.2.259

Khan, U., Dhar, R., & Wertenbroch, K. (2005). A behavioral decision theory perspective on hedonic and utilitarian choice. In S. Ratneshwar & D. G. Mick (Eds.), *Inside consumption: Frontiers of research on consumer motives, goals, and desires* (pp. 144-165). London, U.K.: Routledge.

Kivetz, R., & Simonson, I. (2002). Self-control for the righteous: Toward a theory of precommitment to indulgence. *Journal of Consumer Research, 29*(2), 199-217. doi: Doi 10.1086/341571

Kurbat, M. A., Shevell, S. K., & Rips, L. J. (1998). A year's memories: The calendar effect in autobiographical recall. *Memory & Cognition, 26*(3), 532-552. doi: 10.3758/Bf03201161

LeBoeuf, R. A., Williams, E. F., & Brenner, L. A. (2014). Forceful phantom firsts: Framing experiences as firsts amplifies their influence on judgment. *Journal of Marketing Research, 51*(4), 420-432. doi: 10.1509/Jmr.11.0298

Libby, L. K., & Eibach, R. P. (2002). Looking back in time: Self-concept change affects visual perspective in autobiographical memory. *Journal of Personality and Social Psychology, 82*(2), 167-179. doi: 10.1037//0022-3514.82.2.167

Libby, L. K., & Eibach, R. P. (2011). Self-enhancement or self-coherence? Why people shift visual perspective in mental images of the personal past and future. *Personality and Social Psychology Bulletin, 37*(5), 714-726. doi: 10.1177/0146167211400207

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. Journal of Personality and Social Psychology, 75(1), 5-18. doi: 10.1037//0022-3514.75.1.5

Liu, W., & Aaker, J. (2007). Do you look to the future or focus on today? The impact of life experience on intertemporal decisions. Organizational Behavior and Human Decision Processes, 102(2), 212-225. doi: 10.1016/j.obhdp.2006.02.004

Locke, E. A., & Latham, G. P. (1990). A theory of goal setting and task performance. Englewood Cliffs, NJ: Prentice-Hall.

Markus, H. R., Mullally, P., & Kitayama, S. (1997). Selfways: Diversity in modes of cultural participation. In U. Neisser & D. Jopling (Eds.), The conceptual self in context: Culture, experience, selfunderstanding (pp. 13-61). Cambridge, England: Cambridge University Press.

Marlatt, G. A., & Kaplan, B. E. (1972). Self-initiated attempts to change behavior - a study of new years resolutions. *Psychological Reports, 30*(1), 123-131.

Milkman, K. L. (2012). Unsure what the future will bring? You may overindulge: Uncertainty increases the appeal of wants over shoulds. *Organizational Behavior and Human Decision Processes, 119*(2), 163-176. doi: 10.1016/j.obhdp.2012.07.003

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2008). Harnessing our inner angels and demons what we have learned about want/should conflicts and how that knowledge can help us reduce short-sighted decision making. *Perspectives on Psychological Science, 3*(4), 324-338. doi: 10.1111/j.1745-6924.2008.00083.x

Norcross, J. C., Mrykalo, M. S., & Blagys, M. D. (2002). Auld lang syne: Success predictors, change processes, and self-reported outcomes of new year's resolvers and nonresolvers. *Journal of Clinical Psychology, 58*(4), 397-405. doi: 10.1002/Jclp.1151

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review, 89*(1), 103-124. doi: 10.1257/Aer.89.1.103

Parfit, D. (1984). *Reasons and persons*. Oxford, England: Oxford University Press.

Peetz, J., & Wilson, A. E. (2013). The post-birthday world: Consequences of temporal landmarks for temporal self-appraisal and motivation. *Journal of Personality and Social Psychology, 104*(2), 249-266. doi: 10.1037/A0030477

Polivy, J., & Herman, C. P. (2002). If at first you don't succeed: False hopes of self-change. *American Psychologist, 57*(9), 677-689. doi: 10.1037//0003-066x.57.9.677

Pollack, A. (2005). Maker of diet pill reports success in a trial. *New York Times*, Retrieved May 11, 2013 from http://www.nytimes.com/2005/05/11/business/11drug.html.

Rajagopal, P., & Rha, J.-Y. (2009). The mental accounting of time. *Journal of Economic Psychology, 30*, 772-781.

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty, 19*(1-3), 171-197. doi: 10.1023/A:1007879411489

Robinson, J. A. (1986). Temporal reference systems and autobiographical memory. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 159-188). Cambridge, England: Cambridge University Press.

Rogers, T., & Bazerman, M. H. (2008). Future lock-in: Future implementation increases selection of 'should' choices. *Organizational Behavior and Human Decision Processes, 106*(1), 1-20. doi: 10.1016/j.obhdp.2007.08.001

Rubin, D. C., & Kozin, M. (1984). Vivid memories. *Cognition, 16*(1), 81-95. doi: 10.1016/0010-0277(84)90037-4

Sackett, A. M., Wu, G., White, R. J., & Markel, A. B. (2012). *Harnessing optimism: How eliciting goals improves performance*. Working Paper. University of St. Thomas, Minneapolis.

Schwarz, J. (1997). How to keep up with those new year's resolutions, researchers find commitment is the secret of success. Retrieved May 28, 2013 from http://www.washington.edu/news/1997/12/23/how-to-keep-up-with-those-new-years-resolutions-researchers-find-commitment-is-the-secret-of-success/.

Sela, A., Berger, J. A., & Liu, W. (2009). Variety, vice, and virtue: How assortment size influences option choice. *Journal of Consumer Research, 35*(6), 941-951.

Shai, I., Schwarzfuchs, D., Henkin, Y., Shahar, D. R., Witkow, S., Greenberg, I., . . . Grp, D. (2008). Weight loss with a low-carbohydrate, mediterranean, or low-fat diet.

New England Journal of Medicine, 359(3), 229-241. doi:

10.1056/Nejmoa0708681

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and

cognition in consumer decision making. *Journal of Consumer Research, 26*(3),

278-292. doi: 10.1086/209563

Shoda, Y., Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-

regulatory competences from preschool delay of gratification - identifying

diagnostic conditions. *Developmental Psychology, 26*(6), 978-986. doi:

10.1037//0012-1649.26.6.978

Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory

processes. *Psychological Bulletin, 124*(3), 423-442. doi: 10.1037//0033-

2909.124.3.423

Soman, D. (2001). The mental accounting of sunk time costs: Why time is not like

money. *Journal of Behavioral Decision Making, 14*(3), 169-185. doi:

10.1002/Bdm.370

Soman, D. (2004). Framing, loss aversion, and mental accounting. In D. J. Hoehler & N.

Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 379-

398). Boston, MA: Blackwell Publishing.

Soman, D., & Ahn, H.-K. (2011). Mental accounting and individual welfare. In K.

Gideon (Ed.), *Perspectives on framing* (pp. 65-92). New York, NY: Psychology

Press.

Soster, R. L., Monga, A., & Bearden, W. O. (2010). Tracking costs of time and money: How accounting periods affect mental accounting. *Journal of Consumer Research, 37*, 712–721.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making, 12*(3), 183-206.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge*. New Haven, CT: Yale University Press.

Toure-Tillery, M., & Fishbach, A. (2012). The end justifies the means, but only in the middle. *Journal of Experimental Psychology-General, 141*(3), 570-583. doi: 10.1037/A0025928

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review, 110*(3), 403-421. doi: 10.1037/0033-295x.110.3.403

Tu, Y. P., & Soman, D. (2014). The categorization of time and its impact on task initiation. *Journal of Consumer Research, 41*(3), 810-822. doi: 10.1086/677840

USA.gov. (2013). Popular new year's resolutions. Retrieved January 14, 2013 from http://www.usa.gov/Citizen/Topics/New-Years-Resolutions.shtml.

Wilson, A. E., & Ross, M. (2001). From chump to champ: People's appraisals of their earlier and present selves. *Journal of Personality and Social Psychology, 80*(4), 572-584. doi: 10.1037//0022-3514.80.4.572

Wilson, A. E., & Ross, M. (2003). The identity function of autobiographical memory: Time is on our side. *Memory, 11*(2), 137-149. doi: 10.1080/09658210244000324

Wood, W., Tam, L., & Witt, M. G. (2005). Changing circumstances, disrupting habits. *Journal of Personality and Social Psychology, 88*(6), 918-933. doi: 10.1037/0022-3514.88.6.918

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge, MA: The MIT Press.

Zauberman, G., Levav, J., Diehl, K., & Bhargave, R. (2010). 1995 feels so close yet so far: The effect of event markers on subjective feelings of elapsed time. *Psychological Science, 21*(1), 133-139. doi: 10.1177/0956797609356420

Zhang, Y., Fishbach, A., & Dhar, R. (2007). When thinking beats doing: The role of optimistic expectations in goal-based choice. *Journal of Consumer Research, 34*, 567-578.

Table 1. Ordinary Least Squares Regressions to Predict Daily Google Search Volume for

Various Search Terms (Study 1)

Google search term:	Diet	News	Weather	Laundry	Gardening
	Model 1	Model 2	Model 3	Model 4	Model 5
Generic calendar predictors					
Days since the start of the week (Monday)	-1.63***	-2.09***	0.72***	1.89***	2.23***
	(0.08)	(0.11)	(0.17)	(0.10)	(0.13)
Days since the start of the month	-0.09***	-0.05*	0.09^	1.8e-03	-0.07
	(0.02)	(0.02)	(0.04)	(0.02)	(0.06)
Months since the start of the year	-3.81***	-0.05	0.93	-1.02*	-1.29
	(0.42)	(0.45)	(0.83)	(0.41)	(1.88)
Work calendar predictors					
First workday after a federal holiday	7.40***	-1.76*	0.77	2.89**	0.29
	(0.73)	(0.84)	(0.76)	(0.94)	(1.11)
First workday x Fresh start score of federal holiday	6.78***	-2.19***	2.27*	-0.26	-3.25***
	(0.65)	(0.54)	(0.73)	(0.39)	(0.75)
Fixed effects for each three-month download interval	Yes	Yes	Yes	Yes	Yes
Observations	3,104	3,104	3,104	3,104	3,104
R^2	0.62	0.81	0.53	0.33	0.32

Notes. Model 1 reports the coefficients from an OLS regression predicting the relative Google search volume for "diet" as a function of a given day's proximity to a variety of calendar markers. Models 2-5 predict search volume for the placebo terms "news", "weather", "laundry", and "gardening" respectively, using the same regression specification as Model 1. Standard errors (in parentheses) are clustered at the three-month interval level.

$^\wedge p < 0.10;$ $^* p < 0.05;$ $^{**} p < 0.01;$ $^{***} p < 0.001$

Table 2. Ordinary Least Squares Regressions to Predict Daily Undergraduate Gym

Attendance (Study 2)

Sample:	All undergraduate gym members		Members with birthday information
Regression outcome variable:	Average gym visits per hour	Daily individual indicator[b]	Daily individual indicator[b]
	Model 6	Model 7	Model 8
Generic calendar predictors			
Days since the start of the week (Monday)	-2.12***	-3.5e-03***	-3.5e-03***
	(0.38)	(6.8e-04)	(7.5e-04)
Days since the start of the month	-0.37***	-4.3e-04***	-3.9e-04**
	(0.09)	(1.2e-04)	(1.3e-04)
Months since the start of the year	-0.66**	-9.6e-04**	-7.1e-04*
	(0.21)	(2.8e-04)	(3.0e-04)
Academic calendar predictors			
Months since the start of the semester	-6.97***	-9.8e-03***	-9.7e-03***
	(0.91)	(1.1e-03)	(1.3e-03)
First day after a school break	15.53***	0.02**	0.03**
	(4.41)	(8.3e-03)	(9.4e-03)
Personal calendar predictor			
Months since last birthday			-5.9e-04***
			(1.0e-04)
Controls for school session[a]	Yes	Yes[b]	Yes[b]
Fixed effects for each gym member	No	Yes	Yes
Observations	442	5,265,104	722,362
Number of gym members	11,912	11,912	2,076
R^2	0.67	0.14	0.14

Notes. Models 6–8 report the results from OLS regressions in which the dependent measure is the daily average visits per hour at a university gym (Model 6) and the likelihood that a given person visited the university gym on a given day (Models 7 and 8). Standard errors (in parentheses) are clustered at the date level in Models 7 and 8. Predictor variables include measures of a given day's proximity to a variety of temporal landmarks.

[a] School session control variables include normal school session indicator (during the fall and spring semesters), summer session indicator, final exam period indicator, and days since the exam period starts.

[b] Besides school session control variables, the number of operating hours on each date is included as a control variable.

^ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 3. Summary Statistics for Goal Contracts Created on stickK.com from October 1, 2010, to February 13, 2013, by Goal Category (Study 3)

	Total contracts		Daily contracts			
	Sum	% of all contracts	Mean	SD	Max	Min
Custom goal	28,830	43.64	33.29	19.58	174	3
Health-irrelevant custom goal[a]	15,213	23.03	17.57	11.25	94	1
Health-relevant custom goal[a]	12,976	19.64	15.04	9.25	78	0
Exercise regularly	10,759	16.29	12.42	10.73	140	0
Lose weight	23,823	36.06	27.51	27.14	349	4
Maintain weight	403	0.61	0.47	0.70	4	0
Quit smoking	1,500	2.27	1.73	1.88	24	0
Run a race	747	1.13	0.65	1.10	7	0
All types of goals	66,062	100.00	76.28	55.68	687	16

[a] The data set does not contain sub-category information for all custom goals, but instead for a subset of 28,189 (or 98% of) custom goals.

Table 4. Ordinary Least Squares Regressions to Predict Daily Creation of Commitment

Contracts on stickK.com in Aggregate (Study 3)

Goal category:	All categories		Health-irrelevant custom goals	
Regression outcome variable:	Daily number of contracts	Did individual create a goal? (Y=1, N=0)	Daily number of contracts	Did individual create a goal? (Y=1, N=0)
	Model 9	Model 10[a]	Model 11	Model 12[a]
Generic calendar predictors				
Days since the start of the week (Monday)	-5.73***	-1.2e-04***	-1.12***	-1.0e-04***
	(0.79)	(1.7e-05)	(0.16)	(1.2e-05)
Days since the start of the month	-0.49**	-1.0e-05**	-0.06	-4.2e-06
	(0.18)	(3.2e-06)	(0.04)	(2.8e-06)
Months since the start of the year	-6.09***	-1.2e-04***	-1.24***	-1.0e-04***
	(0.43)	(1.1e-05)	(0.08)	(8.8e-06)
Work calendar predictors				
First workday after a federal holiday	41.96***	8.0e-04**	5.81**	4.9e-04*
	(9.34)	(2.3e-04)	(1.84)	(2.2e-04)
First workday x Fresh start score of federal holiday	67.74***	1.3e-03**	8.97***	6.0e-04*
	(8.60)	(3.7e-04)	(1.70)	(2.3e-04)
Personal calendar predictor				
Months since last birthday		-3.4e-06^		-5.8e-06
		(1.9e-06)		(3.7e-06)
Days since launch	-0.01	-1.8e-07	-0.01*	-6.6e-07
	(0.03)	(5.1e-07)	(5.0e-03)	(4.5e-07)
Days since launch^2	5.5e-05^	9.7e-10^	2.8e-05***	2.1e-09***
	(2.8e-05)	(5.8e-10)	(5.6e-06)	(5.4e-10)
Fixed effects for each stickK user	No	Yes	No	Yes
Observations	866	37,162,658	866	8,694,640
Number of stickK users[b]	43,012	42,913	10,074	10,040
R^2	0.32	2.3e-03	0.35	2.5e-03

Notes. Models 9 and 11 predict the daily number of commitment contracts associated with all types of goals (Model 9) and health-irrelevant custom goals (Model 11). Models 10c and 12 predict the likelihood that a given user created a goal contract on a given day for all types of goals (Model 10) and for health-irrelevant custom goals (Model 12). Standard errors (in parentheses) are clustered at the date level for Models 10 and 12. Across all models, independent variables include measures of a given day's proximity to a variety of temporal

[a] This regression model includes the 99.5% of users whose birthdates were available to us.

[b] This represents the number of stickK users who created at least one commitment contract in a corresponding goal category and thus were included in each corresponding regression model.

[c] We only include regression results for the 42,913 users with a known birthdate because these users account for more than 99.5% of all users in our data set. When we predict the likelihood of creating a commitment contract on a given day as a function of the aforementioned predictors (with the exception of months since last birthday) for all 43,012 users in our data set, the regression results we obtain are virtually identical.

^ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 5. Ordinary Least Squares Regressions to Predict Daily Creation of Commitment

Contracts on stickK.com by Goal Category (Study 3)

Regression outcome variable:	Daily number of contracts								
Goal category:	Career	Education and knowledge	Money and finance	Health-relevant custom goals	Regular exercise	Weight loss	Weight maintenance	Smoking cessation	Running a race
	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19	Model 20	Model 21
Generic calendar predictors									
Days since the start of the week (Monday)	-0.59***	-0.25***	-0.08*	-0.95***	-1.03***	-2.30***	-0.04**	-0.17***	-0.09***
	(0.05)	(0.06)	(0.03)	(0.13)	(0.16)	(0.41)	(0.01)	(0.03)	(0.02)
Days since the start of the month	-0.01	-0.02	-0.02*	-0.06*	-0.07*	-0.29**	-2.1e-03	-0.01	8.1e-05
	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.09)	(2.7e-03)	(0.01)	(4.0e-03)
Months since the start of the year	-0.33***	-0.30***	-0.13***	-1.05***	-1.08***	-2.41***	-0.02**	-0.12***	-0.07***
	(0.03)	(0.03)	(0.02)	(0.07)	(0.09)	(0.22)	(6.5e-03)	(0.02)	(9.8e-03)
Work calendar predictors									
First workday after a federal holiday	0.16	1.93**	0.94*	6.77***	7.37***	20.10***	0.05	1.16**	0.55*
	(0.64)	(0.73)	(0.41)	(1.54)	(1.87)	(4.80)	(0.14)	(0.35)	(0.21)
First workday x Fresh start score of federal holiday	1.85**	0.63	2.21***	7.48***	13.22***	36.64***	9.6e-03	1.42***	0.34^
	(0.59)	(0.67)	(0.37)	(1.42)	(1.72)	(4.42)	(0.13)	(0.33)	(0.20)
Days since launch	-1.3e-03	-1.3e-03	9.0e-05	9.2e-03***	-3.8e-03	4.8e-03	6.6e-04^	6.9e-04	-6.4e-04
	(1.7e-03)	(2.0e-03)	(1.1e-03)	(1.0e-03)	(5.1e-03)	(0.01)	(3.8e-04)	(9.6e-04)	(5.8e-04)
Days since launch^2	7.0e-06***	5.5e-06*	1.2e-06	1.9e-06	9.4e-06	1.9e-06	-6.7e-07	-1.4e-07	8.3e-07
	(1.9e-06)	(2.2e-06)	(1.2e-06)	(1.0e-03)	(5.7e-06)	(1.5e-05)	(4.3e-07)	(1.1e-06)	(6.4e-07)
Observations	866	866	866	866	866	866	866	866	866
Number of stickK users[a]	3,068	2,944	1,399	8,493	9,695	20,273	327	1,329	700
R^2	0.31	0.16	0.13	0.33	0.27	0.25	0.03	0.14	0.10

Notes. Models 13–21 predict the daily number of commitment contracts associated with each of the three most popular health-irrelevant custom goals (Models 13–15), all health-irrelevant custom goals combined (Model 16), as well as each of the five standard goals (Models 17–21). Across all models, independent variables include measures of a given day's proximity to a variety of temporal landmarks.

[a] This represents the number of stickK users who created at least one commitment contract in a corresponding goal category and thus were included in each corresponding regression model.

^ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

68

Figure 1. Changes in the Fitted Daily Search Volume for the Term "Diet" as a Function

of the Date and Its Proximity to a Variety of Temporal Landmarks (Study 1)

Fitted change in the daily Google search volume for the word "diet"
(search volume scaled from 0-100)
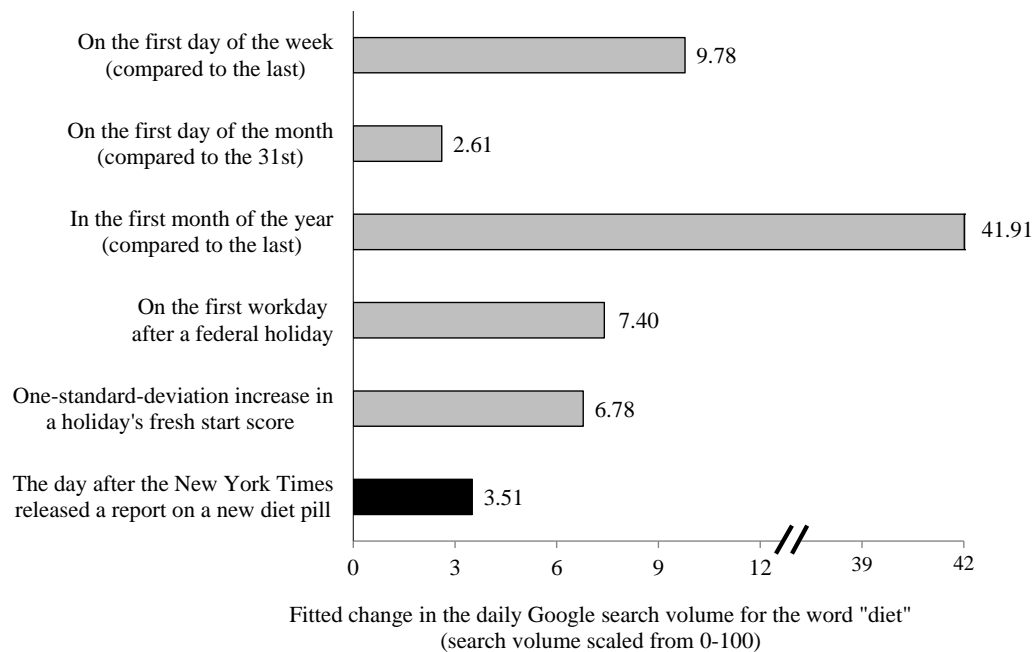
Note. These effects are compared with the effect of the *New York Times* releasing a report about a promising new diet pill (Pollack 2005) on searches for the term "diet."

Figure 2. Changes in the Fitted Probability of Going to the Gym as a Function of the

Date and Its Proximity to a Variety of Temporal Landmarks (Study 2)
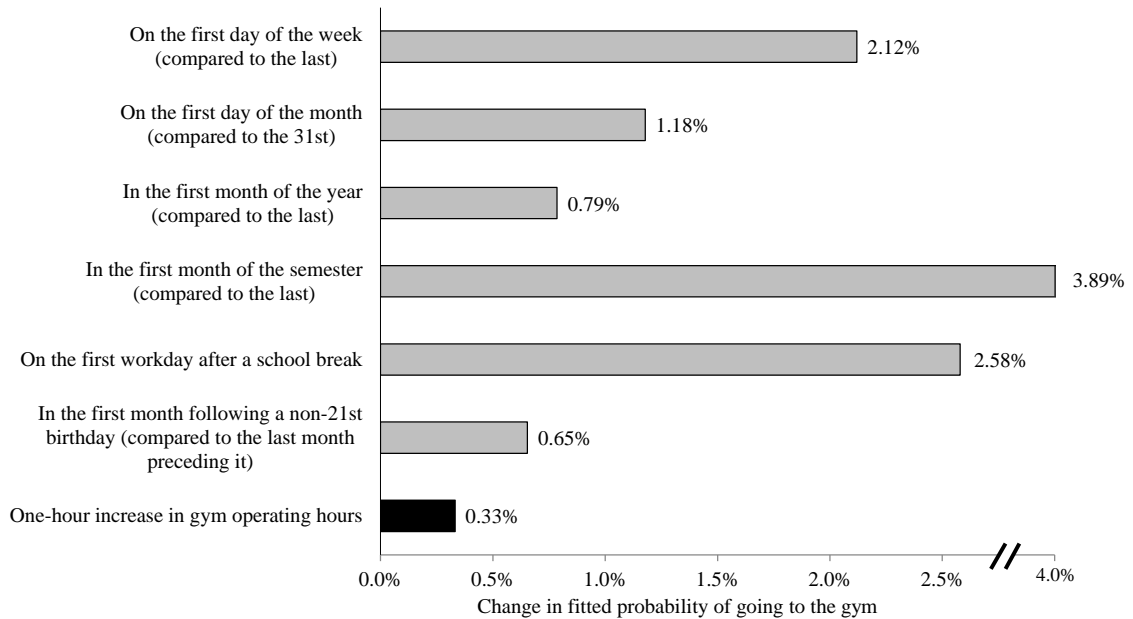
Note. These effects are compared with the effect of a one-hour increase in the gym's
operating hours on the likelihood of going to the gym.

Figure 3. Changes in the Fitted Probability of Creating a Commitment Contract as a

Function of the Date and Its Proximity to a Variety of Temporal Landmarks (Study 3)
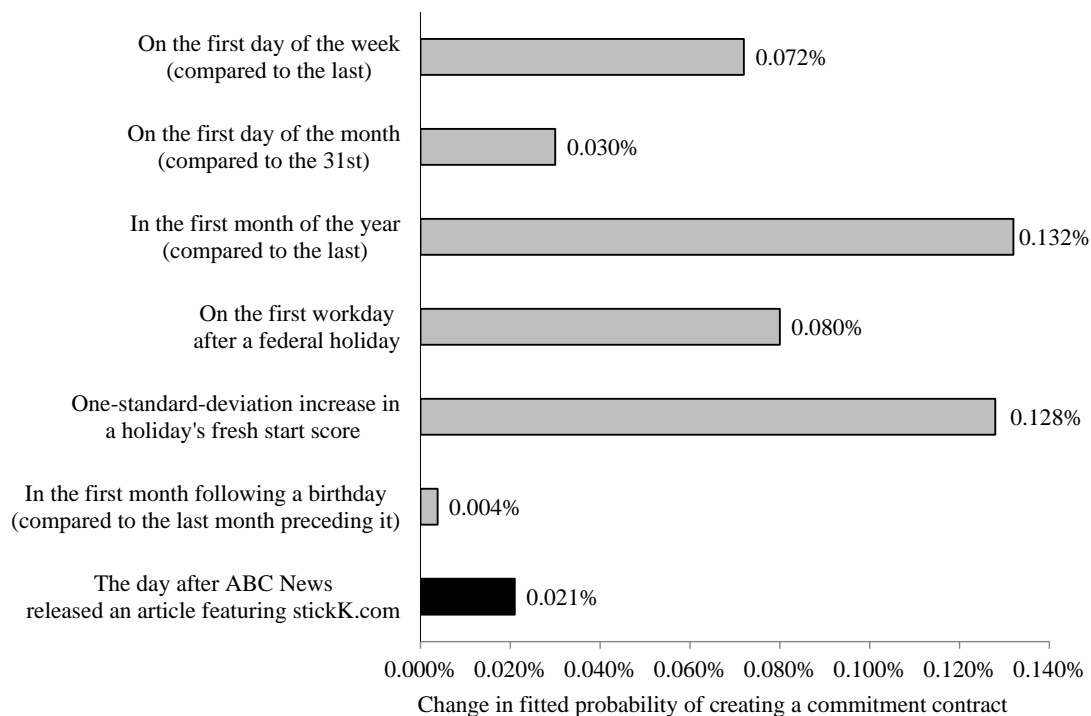
Note. These effects are compared with the effect of *ABC News* releasing an article featuring stickK (Farnham 2012) on the likelihood of creating a commitment contract.

Appendix A. Description of Data from Google Insights for Search (Analyzed in Study 1)

Google Insights for Search provides data on the relative search volume for a given term on Google based on the raw number of searches for that term. The data provided by Google has the following features, which are described on the Google Insights for Search website:

1.   It excludes repeated queries: "Our system also eliminates repeated queries from a single user over a short period of time, so that the level of interest isn't artificially impacted by these types of queries."[16]

2.   Data is normalized: "…All the results in Google Insights for Search are normalized, which means that we've divided the sets of data by a common variable to cancel out the variable's effect on the data. Doing so allows the underlying characteristics of the data sets to be compared. If we didn't normalize the results and displayed the absolute rankings instead, data from regions generating the most search volume would always be ranked high."[17] "Google Insights for Search analyzes a portion of Google web searches to compute how many searches have been done for the terms you've entered, relative to the total number of searches done on Google over time."[18]

3.   Data is scaled: "The data is displayed on a scale of 0 to 100. To arrive at those values, we first normalize the data. After normalization, we divide each point on the graph by the highest value and then multiply by 100. Different plots are not comparable unless they share the same original highest value before scaling."[19] Note that we downloaded data in three-month intervals (the longest interval possible at the daily level), so our data is scaled at this level.

4.   Data values below a certain threshold are recorded as zeros: "…Insights designates a certain threshold of traffic for search terms, so that those with low volume won't appear."[20]

---

[16] How is the data derived. (n.d.). Retrieved July, 2012 from
http://support.google.com/insights/bin/answer.py?hl=en&answer=92768&topic=13975&ctx=topic
[17] Is the data normalized. (n.d.) Retrieved July, 2012 from
http://support.google.com/insights/bin/answer.py?hl=en&answer=87284&topic=13975&ctx=topic
[18] How is the data derived. (n.d.). Retrieved July, 2012 from
http://support.google.com/insights/bin/answer.py?hl=en&answer=92768&topic=13975&ctx=topic
[19] How is the data scaled. (n.d.). Retrieved July, 2012 from
http://support.google.com/insights/bin/answer.py?hl=en&answer=87282&topic=13975&ctx=topic.
[20] How is the data derived. (n.d.). Retrieved July, 2012 from
http://support.google.com/insights/bin/answer.py?hl=en&answer=92768&topic=13975&ctx=topic

Appendix B. Survey Question Used to Generate Holidays' Fresh Start Ratings

Question

Think about the following days in a typical man's life. To what extent would each day (or the day just after it) feel like a fresh start? (1 = Not at all; 7=Very much)

List of holidays

- New Year's Day (January 1st)
- Martin Luther King Day (the third Monday of January; January 16th in 2012)
- Groundhog Day (February 2nd)
- Valentine's Day (February 14th)
- President's Day (the third Monday of February; February 20th in 2012)
- The day when Daylight Saving Time begins (the Second Sunday in March; March 11th in 2012)
- St. Patrick's Day (March 17th)
- April Fool's Day (April 1st)
- Tax Day (usually April 15th; April 16th in 2012)
- Earth Day (April 22nd)
- National Secretary's Day (the Wednesday of the last full week of April; April 25th in 2012)
- Mother's Day (the second Sunday of May; May 13th in 2012)
- Memorial Day (the last Monday of May; May 28th in 2012)
- Flag Day (June 14th)
- Father's Day (June 17th in 2012)
- Independence Day (July 4th in 2012)
- Labor Day (the first Monday of September; September 3rd in 2012)
- Patriot Day or 9/11 (Sep 11th)
- Columbus Day (the second Monday of October; October 8th in 2012)
- Halloween (October 31st)
- The day when Daylight Saving Time ends (the first Sunday in November; November 4th in 2012)
- Election Day (the first Tuesday after the first Monday in November; November 6th in 2012)
- Veteran's Day (November 11th)
- Thanksgiving Day (the fourth Thursday of November; November 22nd in 2012)
- Black Friday (the day after Thanksgiving; November 23rd in 2012)
- Christmas Day (December 25th)

Appendix C. Additional Regression Results with Google Search Data (Study 1)

Google search term:	**Diet**	**News**	**Weather**	**Laundry**	**Gardening**
	Model A1	Model A2	Model A3	Model A4	Model A5
Generic calendar predictors					
Tuesday	-3.82***	1.15***	-2.75***	-5.40***	-5.02***
	(0.33)	(0.27)	(0.41)	(0.39)	(0.55)
Wednesday	-5.10***	2.69***	-1.26*	-6.82***	-7.28***
	(0.37)	(0.36)	(0.50)	(0.49)	(0.47)
Thursday	-8.36***	2.90***	0.89	-7.33***	-8.11***
	(0.41)	(0.39)	(0.63)	(0.37)	(0.46)
Friday	-13.82***	1.75***	3.20**	-7.84***	-7.87***
	(0.40)	(0.40)	(0.83)	(0.36)	(0.43)
Saturday	-13.50***	-10.70***	3.39**	3.78***	6.02***
	(0.42)	(0.59)	(0.95)	(0.55)	(0.64)
Sunday	-5.82***	-11.18***	1.24^	11.74***	13.55***
	(0.38)	(0.66)	(0.66)	(0.74)	(0.80)
Days since the start of the month	-0.09***	-0.04^	0.09^	-1.4e-03	-0.07
	(0.02)	(0.02)	(0.04)	(0.02)	(0.06)
Months since the start of the year	-3.78***	-0.10	0.94	-0.95*	-1.21
	(0.42)	(0.45)	(0.83)	(0.40)	(1.88)
Work calendar predictors					
First workday after a federal holiday	7.21***	-0.59	1.70^	2.07*	-0.99
	(0.69)	(0.84)	(0.85)	(0.94)	(1.01)
First workday x Fresh start score of federal holiday	6.99***	-2.23**	1.65	-0.41	-3.31***
	(0.63)	(0.68)	(1.07)	(0.54)	(0.82)
Fixed effects for each three-month download interval	Yes	Yes	Yes	Yes	Yes
Observations	3,104	3,104	3,104	3,104	3,104
R^2	0.69	0.86	0.54	0.58	0.46

Notes. Models A1-A5 report coefficients from OLS regressions predicting the relative Google search volume for the term "diet" and the placebo terms "news," "weather," "laundry," and "gardening" respectively as a function of a given day's proximity to a variety of breaking points. The models here rely on identical specifications to those presented in Table 1, Models 1-5 in the main manuscript with one exception: the continuous measure days since the start of the week is replaced with six indicator variables – one for each day of the week from Tuesday to Sunday (with Monday omitted). Standard errors are clustered at the three-month interval level; these intervals correspond to the time periods over which Google's data was automatically scaled (see Appendix A).

^ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Appendix D. Additional Regression Results with Gym Attendance Data (Study 2)

Sample:	All undergraduate gym members		Members with birthday information
Regression outcome variable:	Average gym visits per hour	Daily individual indicator[b]	Daily individual indicator[b]
	Model A6	Model A7	Model A8
Generic calendar predictors			
Tuesday	-3.40	-3.1e-03	-5.9e-03
	(2.77)	(3.7e-03)	(4.0e-03)
Wednesday	-7.68**	-0.01*	-0.01*
	(2.75)	(3.7e-03)	(4.0e-03)
Thursday	-9.29**	-0.01**	-0.01**
	(2.78)	(3.6e-03)	(3.9e-03)
Friday	-10.96***	-0.01**	-0.01***
	(2.80)	(4.0e-03)	(4.2e-03)
Saturday	-11.09***	-0.02***	-0.02**
	(2.81)	(5.3e-03)	(5.9e-03)
Sunday	-13.86***	-0.02***	-0.02***
	(2.80)	(4.8e-03)	(5.3e-03)
Days since the start of the month	-0.38***	-2.4e-04*	-3.9e-04**
	(0.09)	(1.1e-04)	(1.3e-04)
Months since the start of the year	-0.66**	-1.1e-03***	-7.1e-04*
	(0.22)	(2.8e-04)	(3.0e-04)
Academic calendar predictors			
Months since the start of the semester	-7.02***	-5.4e-03***	-9.7e-03***
	(0.91)	(8.5e-04)	(1.3e-03)
First day after a school break	15.61**	0.03**	0.03**
	(4.47)	(8.4e-03)	(9.3e-03)
Personal calendar predictor			
Months since last birthday			-5.9e-04***
			(1.0e-04)
Controls for school session[a]	Yes	Yes[b]	Yes[b]
Fixed effects for each gym member	No	Yes	Yes
Observations	442	5,265,104	722,362
Number of gym members	11,912	11,912	2,076
R^2	0.67	0.14	0.14

Notes. Models A6-A8 report coefficients from OLS regressions in which the dependent measure is the daily average hourly attendance at the university gym in our study (Model A6) and the likelihood that a given person visited the university gym on a given day (Models A7-A8). The models here rely on identical specifications to those presented in Table 2, Models 6-8 in the main manuscript with one exception: the continuous measure days since the start of the week is replaced with six indicator variables – one for each day of the week from Tuesday to Sunday (with Monday omitted). Standard errors are clustered at the date level in Models 7 and 8.

[a] School session control variables include normal school session indicator (during the fall and spring semesters), summer session indicator, final exam period indicator, and days since the exam period starts.

[b] Besides school session control variables, the number of operating hours on each date is included as a control variable.

$^p < 0.10$; $* p < 0.05$; $** p < 0.01$; $*** p < 0.001$

Appendix E. Additional Regression Results with stickK Commitment Contract Data (Study 3)

Goal category:	All categories		Health-irrelevant custom goals	
Regression outcome variable:	Daily number of contracts	Did individual create a goal? (Y=1, N=0)	Daily number of contracts	Did individual create a goal? (Y=1, N=0)
	Model A9	Model A10[a]	Model A11	Model A12[a]
Generic calendar predictors				
Tuesday	-10.70^	-2.0e-04	-2.64*	-2.1e-04*
	(5.89)	(1.5e-04)	(1.16)	(1.0e-04)
Wednesday	-16.28**	-3.0e-04*	-3.75**	-3.0e-04**
	(5.83)	(1.4e-04)	(1.15)	(9.3e-05)
Thursday	-23.45***	-4.4e-04***	-3.85**	-3.1e-04**
	(5.84)	(1.1e-04)	(1.15)	(9.1e-05)
Friday	-31.38***	-6.3e-04***	-5.62***	-4.9e-04***
	(5.83)	(1.1e-04)	(1.15)	(8.3e-05)
Saturday	-43.88***	-8.7e-04***	-9.34***	-7.7e-04***
	(5.83)	(1.1e-04)	(1.15)	(8.0e-05)
Sunday	-26.39***	-5.7e-04***	-5.39***	-5.0e-04***
	(5.83)	(1.2e-04)	(1.15)	(8.9e-05)
Days since the start of the month	-0.50**	-1.0e-05**	-0.06^	-4.3e-06
	(0.18)	(3.2e-06)	(0.03)	(2.8e-06)
Months since the start of the year	-6.08***	-1.2e-04***	-1.24***	-1.0e-04***
	(0.43)	(1.1e-05)	(0.08)	(8.7e-06)
Work calendar predictors				
First workday after a federal holiday	41.45***	7.9e-04**	5.94**	5.0e-04*
	(9.45)	(2.4e-04)	(1.86)	(2.2e-04)
First workday x Fresh start score of federal holiday	67.06***	1.3e-03**	8.81***	5.9e-04*
	(8.60)	(3.7e-04)	(1.69)	(2.3e-04)
Personal calendar predictor				
Months since last birthday		-3.4e-06^		-5.8e-06
		(1.9e-06)		(3.7e-06)
Days since launch	-0.01	-1.8e-07	-0.01*	-6.6e-07
	(0.03)	(5.0e-07)	(4.9e-03)	(4.4e-07)
Days since launch^2	5.5e-05^	9.7e-10^	2.8e-05***	2.1e-09***
	(2.8e-05)	(5.8e-10)	(5.5e-06)	(5.3e-10)
Fixed effects for each stickK user	No	Yes	No	Yes
Observations	866	37,162,658	866	8,694,640
Number of stickK users[b]	43,012	42,913	10,074	10,040
R^2	0.33	2.3e-03	0.37	2.5e-03

Notes. Models A9 and A11 predict the daily number of commitment contracts associated with all types of goals (Model A9) and health-irrelevant custom goals (Model A11). Models A10 and A12 predict the likelihood that a given user created a goal contract on a given day for all types of goals (Model A10) and for healthirrelevant custom goals (Model A12). The models here rely on identical specifications to those presented in Table 4, Models 9-12 in the main manuscript with one exception: the continuous measure days since the start of the week is replaced with six indicator variables – one for each day of the week from Tuesday to Sunday (with Monday omitted). Standard errors are clustered at the date level for Models A10 and A12.

[a] This regression model includes the 99.5% of users whose birthdates were available to us.

[b] This represents the number of stickK users who created at least one commitment contract in a corresponding goal category and thus were included in each corresponding regression model.

^ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Appendix F. Surveys Used to Generate Placebo Search Terms

<u>Instructions on listing activities that people tend to put off</u>

There are some activities that you tend to engage in right away when the inclination strikes you (e.g., getting food when you feel hungry, watching TV when you want to learn the latest news). There are other activities that you tend to put off because you think you will have more time and energy later than you have now.

Please list three activities that you have the tendency to put off doing until a future date when you think you will have more time and energy.

<u>Instructions on listing "new" activities that people do not engage in regularly</u>

There are some activities that you engage in on a daily basis (e.g., eating, reading newspapers). There are also activities that you do not engage in regularly but may consider doing when you want to start something new.

Please list three "new" activities (i.e., not activities of the type that you engage in on a daily basis) that you have never engaged in before but would consider pursuing in the future.

<u>Instructions on selecting activities that are not aspirational</u>

On the previous page, you listed several activities that are reprinted below.

[Insert here activities that participants listed]

Please read the following definition closely:

Aspirational activities are activities that you are aspired to do. They could be any activities that help you achieve your goals or activities that you set a goal to engage in.

Please check the box in front of any activity you do **<u>NOT</u>** think is aspirational.

Appendix G. Raw Plots of Representative Year-long Periods in Studies 1-3

Panel A. Daily Google search volume for "diet" by date for 2011 (Study 1)
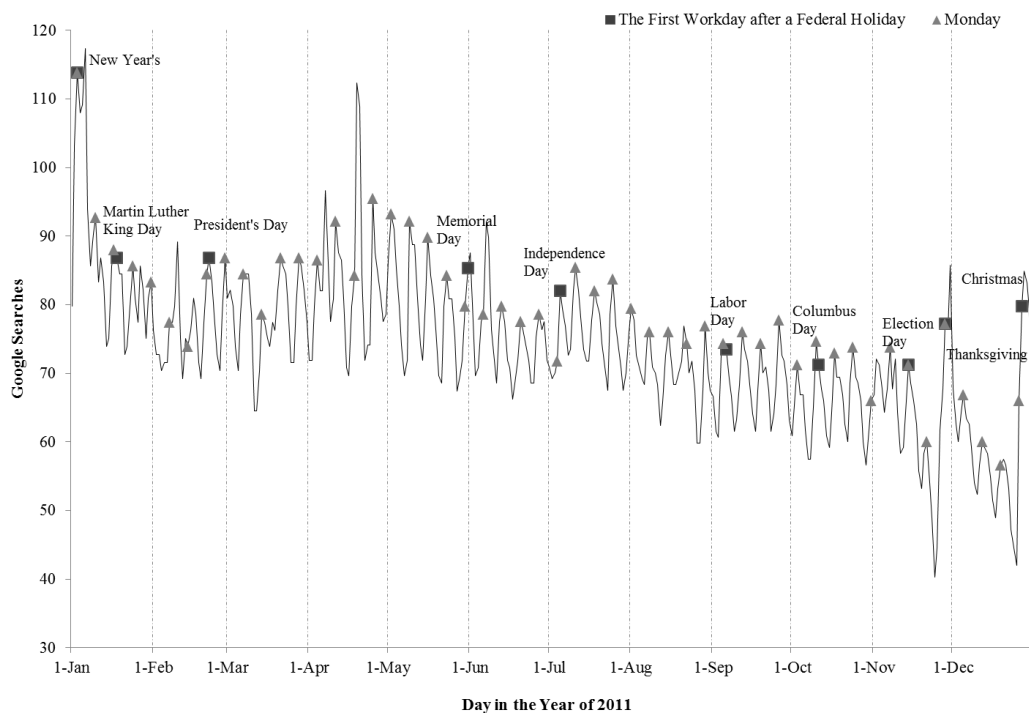
Panel B. The average hourly gym attendance by date of the 2010-2011 calendar year (Study 2)
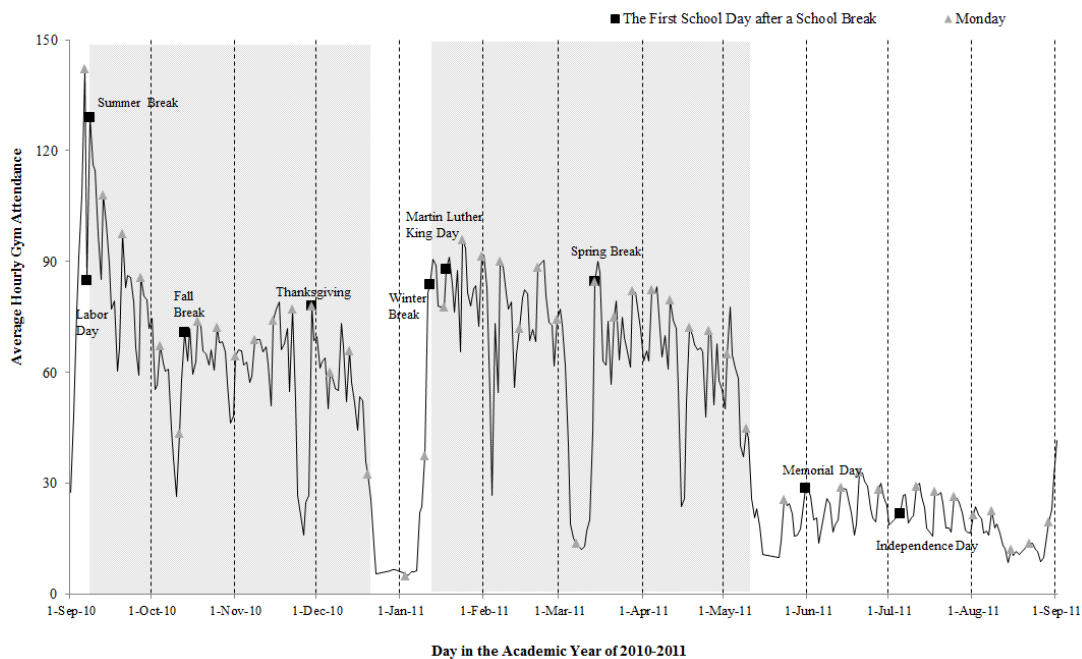
Panel C. The average daily commitment contracts created on stickK.com by date for 2011 (Study 3)
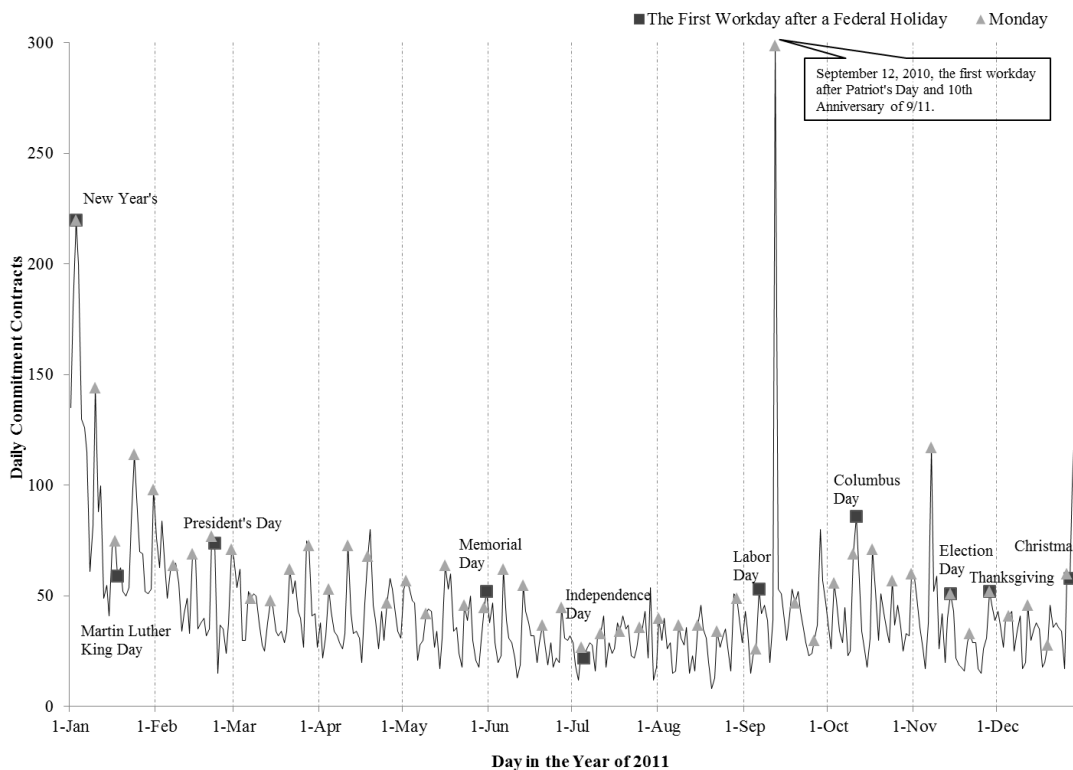
Notes. Panel A depicts daily searches for the term "diet" on Google from January 1, 2011 to December 31, 2011. The raw data from Google Insights for Search has been rescaled based on search volumes between April 1, 2004 and June 30, 2004 so it is comparable across all three-month period intervals. Dashed lines correspond to the first day of each month. Each federal holiday is labeled on the first workday following the holiday.

Panel B depicts average hourly gym attendance (equal to the total number of gym visits on a given day divided by the number of hours the gym was open on that day) over the 2010-2011 academic year. Dashed lines correspond to the first day of each month. Days on a grey background arose during while school was in regular session (during the fall or spring semester). Each school break listed on the university's academic calendar is labeled on the first school day following the relevant school break.

Panel C depicts daily commitment contracts on stickK.com from January 1, 2011 to December 31, 2011. Dashed lines correspond to the first day of each month. Each federal holiday is labeled on the first workday following the relevant holiday.

CHAPTER 2

PUT YOUR IMPERFECTIONS BEHIND YOU: WHY AND HOW

MEANINGFUL TEMPORAL LANDMARKS MOTIVATE GOAL PURSUIT

Hengchen Dai, Katherine L. Milkman and Jason Riis

ABSTRACT

Pursuing our goals is challenging and we can't always muster the requisite motivation. Across five laboratory studies, we explore when people naturally experience enhanced motivation to take actions that facilitate goal achievement and why certain days are more likely to spur goal pursuit than others. We provide causal evidence that the arrival or emphasis of a meaningful temporal landmark increases people's intentions to engage in goal pursuit. In addition, we propose and show that people's strengthened motivation to pursue their aspirations following meaningful temporal landmarks originates in part from the greater psychological disassociation these landmarks induce from a person's past, imperfect self.

Keywords: temporal landmarks; multiple selves; mental accounting; motivation; goals

Introduction

Pursuing our goals is challenging and we can't always muster the requisite motivation. On some days we are inspired to push hard: visiting the gym, ordering a salad, hitting the books; on others, we falter and give into unwanted temptations. But what explains this? In this paper, we consider *when* people naturally experience enhanced motivation to act on their goals and *why* motivation is greater on certain days than others. Addressing these questions can both advance our understanding of naturally-arising points in time that spur goal-directed behaviors and suggest when it may be most effective to provide people with encouragement and tools known to facilitate successful goal pursuit.

We propose that temporal landmarks stimulate goal pursuit. Temporal landmarks are days that "stand in marked contrast to the seemingly unending stream of trivial and ordinary occurrences" (Shum, 1998, p.423) in people's lives. They include transition points on social timetables (e.g., holidays, the start of a new week/month/year/semester; Robinson, 1986). They also included personal life events, such as first experiences (e.g., a first date), developmental milestones (e.g., a wedding), and recurring occasions of significance (e.g., a wedding anniversary; Shum, 1998). Temporal landmarks demarcate adjacent time periods—on socially-shared calendars or on an individual's personal life timeline—and help structure our memories and experiences (Robinson, 1986; Shum, 1998).

We theorize that by demarcating and disrupting the continuous flow of time, temporal landmarks can alter our self-evaluations and open new "mental accounts",

thereby stimulating several psychological processes that facilitate goal pursuit. In terms of altering self-evaluations, research has shown that intervening temporal landmarks weaken the psychological connection between one's past, current, and future selves (Bartels & Rips, 2010; Peetz & Wilson, 2013, 2014). People also tend to attribute negative traits and failures to their disconnected past self to maintain a positive image of their current self (Wilson & Ross, 2001, 2003; Ross & Wilson, 2000). We therefore postulate that the arrival of temporal landmarks creates psychological separation from people's past selves and failures, helping people relegate their missteps to the past and elevating their current self-image.

We argue that feeling separated from past failures and imperfections (and the altered self-evaluation that accompanies it) should motivate goal pursuit for several reasons. First, feeling disconnected from past failures may boost people's self-efficacy, or their belief in their ability to carry out their plans and reach their goals (Bandura, 1997; Libby & Eibach, 2002). Since people exert more effort when they experience higher self-efficacy (Bandura, 1977), increased self-efficacy caused by the relegation of past failures to a preceding period should motivate goal pursuit. Further, feeling optimistic and untarnished by the looming specter of past failures may motivate people to behave consistently with their new, positive self-view (Festinger, 1962; Cialdini, 2007). Finally, any new, goal-incongruent behaviors (e.g., eating a cheeseburger instead of dieting) may loom particularly large, when older failures of the past self become less accessible following a temporal landmark. In prospect theory terms, eating a cheeseburger may now feel like a large, initial loss relative to a reset reference point (where the utility curve for

losses is steepest), rather than another small additional loss added to many other accumulated failures (where the utility curve for losses has flattened; Colby & Chapman, 2013; Heath, Larrick, & Wu, 1999; Soman & Cheema, 2004).

All of these theories lead to our central prediction that temporal landmarks will facilitate goal pursuit. However, we expect some temporal landmarks to be more meaningful and stand out more starkly than others as barriers separating our past and current selves. But what makes a landmark more significant and more likely to segregate time into distinct mental accounting periods? Shum (1998) contends that temporal landmarks carry greater importance and meaning when they resonate with our cultural, occupational or religious identity. For example, 36th birthdays should carry more weight for Chinese people than for Americans since they mark the beginning of a new 12-year Chinese zodiac cycle. First experiences (e.g., moving to a new city for the first time) should also be more momentous than repeated events of the same nature (e.g., moving to a new city for the ninth time; Shum, 1998). We propose that the motivation to tackle goals stimulated by temporal landmarks will be magnified when a landmark's meaningfulness is greater or simply made more salient.

Although recent field research has provided some *correlational* evidence that goal-related activities (e.g., dieting, exercising) increase following temporal landmarks (e.g., the start of the week/month/year, holidays, and birthdays; Dai, Milkman, & Riis, 2014), this paper provides a critical first test of whether temporal landmarks *causally* spur goal pursuit. Specifically, we test whether meaningful temporal landmarks—dates imbued with meaning due to their identity-relevance or rarity—are stronger instigators of

goal pursuit than (a) ordinary days and (b) objectively identical but psychologically less meaningful landmarks. Further, we explore one psychological mechanism that we hypothesize contributes to the link between meaningful landmarks and goal pursuit: psychological segregation from past failures.

Study 1

Study 1 provides an initial test of whether individuals are more motivated to pursue health goals on days they perceive as more meaningful.

Method

We recruited participants through Amazon's Mechanical Turk (an online labor market) to take a short survey. Participants first responded to a series of questions about their activities over the last six months, including one inquiry about whether or not they had dieted or considered dieting at any point in the last six months. Unbeknownst to participants, only those who reported they had thought about dieting in the last six months proceeded to our study; the others were directed to an unrelated study of equal length. We aimed for a sample of 50 participants per experimental condition. Based on pretests, we estimated that approximately 80% of participants who entered our survey would have considered dieting in the past six months. Thus, we invited 200 people to participate in our survey. Data collection ceased when total sample size reached our target number. Complete study materials for all studies are available in my supplemental materials.

Our final study sample included 152 participants (59.2% female, $M_{age} = 33$). Participants were provided with a list of 94 days that might arise in their lives, ranging

from the relatively mundane (e.g., a day in the middle of a year; a day they purchased a

new set of dishware) to the personally meaningful (e.g., the first day after a major

birthday; a day they purchased a new car). These days were divided into six unique

categories, including:

- Birthdays (e.g., the 26st, 30th, 34th, 47th) ($N_{birthday\_stimuli} = 24$)

- Calendar dates (e.g., the first day or middle day of a month) ($N_{calendar\_date\_stimuli} = 15$)

- Annual holidays (e.g., Groundhog Day, Thanksgiving Day) ($N_{holiday\_stimuli} = 26$)

- Days when a new person arrives in the participant's life (e.g., a new neighbor, a

  newborn niece) ($N_{new\_person\_stimuli} = 10$)

- Days associated with obtaining a new possession (e.g., a new outfit, a new home)

  ($N_{new\_possession\_stimuli} = 11$)

- Days associated with a job-related change (e.g., an office move, the initiation of a

  new project) ($N_{job\_change\_stimuli} = 8$)

Participants were randomly assigned to one of three experimental conditions.

Participants in the *motivation condition* were asked to rate how likely they would be to

start a diet on or just after each of these 94 days on a seven-point scale (1 = not at all

likely; 7 = very likely). Participants in the *meaningfulness condition* rated the extent to

which each day or the day just after it felt meaningful to them on a seven-point scale (1 =

not at all; 7=very much). Following Baumeister, Vohs, Aaker, and Garbinsky (2013), we

allowed participants in our survey to define meaningfulness as they saw fit rather than

providing a definition. We were primarily interested in the relationship between the

perceived meaningfulness of a given day and participants' motivation to begin a diet after

that day. However, past research suggests that life transitions (e.g., becoming a parent, going to college) may shift an individual's surrounding environment and daily routines, and produce changes in habits for reasons other than those of interest here (Wood, Tam, & Witt, 2005). To address this alternative explanation, we included a third experimental condition: a *routine change condition*. We asked participants randomly assigned to this condition to rate the extent to which their normal daily routines would differ before versus after each of the days in question (1= not at all to 7 = very much).

Results

For each of the 94 days in our study, we averaged responses from participants in a given experimental condition and created three composite scores for the day in question: average *likelihood of beginning a diet*, average *meaningfulness*, and average *anticipated routine change*. We thus created 282 composite scores in total (282 = 94 days x 3 conditions). This allowed us to analyze the relationship between participants' average *likelihood of beginning a diet* on a given day, that day's *meaningfulness*, and that day's *anticipated routine change* across the 94 days studied. As Figure 1 illustrates, there was a significant and positive correlation (correlation coefficient = 0.17, $p < .01$) between how meaningful a given day felt to participants (*Max* = 6.56, *Min* = 1.50, *M* = 3.73, *SD* = 1.21) and how likely participants believed they would be to begin a diet on or immediately following the day in question (*Max* = 5.45, *Min* = 1.90, *M* = 2.97, *SD* = 0.70). Importantly, when we predicted the average *likelihood of beginning a diet* on a given day with its perceived *meaningfulness* using an ordinary least squares regression

controlling for *anticipated routine change*, the relationship between meaningfulness ratings and motivation to begin a diet became even stronger ($\beta = 0.33, p < .01$).

Discussion

This correlational study shows that people feel more motivated to diet following days they perceive as more meaningful. Subsequent experiments will (1) causally test our hypothesis that temporal landmarks, particularly more meaningful landmarks, motivate goal pursuit and (2) more completely rule out routine change as an alternative explanation for our findings by holding anticipated routine change constant. We will also test our proposed mechanism that meaningful temporal landmarks motivate goal pursuit by separating the present self from past imperfections.

<div align="center">

Study 2

</div>

Across two experiments examining real decisions, we next provide a causal test of the hypothesis that temporal landmarks increase people's motivation to tackle their goals. Specifically, we investigate whether people are more likely to choose to receive a goal reminder on a given date when it is highlighted as corresponding to a meaningful temporal landmark.

Study 2A

Method. In early March of 2014, participants interested in learning how to more effectively tackle their goals were recruited through Amazon's Mechanical Turk to take part in a short survey. Participants were first instructed to describe a personal goal that they planned to begin pursuing in April 2014 and then were asked to list one thing they planned to do to facilitate their pursuit of this goal. Next, we offered to send participants

an email reminder in late March that would describe their goal, their plan for

accomplishing it, and a message they could customize. Only participants who signed up

to receive a reminder went on to complete our survey; they comprised our actual study

sample. Participants who chose not to receive a reminder exited our survey. We aimed for

a sample of 100 participants per experimental condition. Based on pretests, we estimated

that approximately 50% of participants who entered our survey would sign-up to receive

a reminder. Thus, we invited 400 people to participate in our survey. Data collection

ceased when total sample size reached our target number. Our final study sample

included 165 participants (61% female, one unspecified; M_{age} = 32, one unspecified) who

signed up to receive a reminder.

At this stage in the study, participants were randomly assigned to one of two

experimental conditions: the *meaningful landmark condition* or the *control condition*. In

both conditions, participants chose when to receive their personalized reminder from a

list of seven consecutive dates ranging from March 18, 2014 (Tuesday) to March 24,

2014 (Monday) with the day of the week indicated in parentheses following the date. Our

subtle manipulation involved highlighting the meaningfulness of March 20, 2014 as the

start of spring (or the Spring Equinox). Specifically, in the *meaningful landmark

condition*, the date – March 20, 2014 – was followed by a description: "(Thursday; The

First Day of Spring 2014)." In the *control condition*, the description following March 20,

2014 framed it as an ordinary day: "(Thursday; The Third Thursday in March 2014)."

After choosing a date from the list, participants were offered the opportunity to customize

the text of their email reminder. Then they rated to what extent March 20, 2014 felt meaningful to them.

Finally, participants provided their email address so they could receive our promised goal reminder.

Results. Our manipulation check confirmed that March 20, 2014 felt more meaningful when it was described as the first day of spring ($M = 3.68$, $SD = 2.02$) than when it was described as the third Thursday in March ($M = 2.45$, $SD = 1.74$, $t(163) = 4.20$, $p < .0001$). Our dependent measure was whether or not participants chose to receive a reminder on March 20, 2014. We predicted that participants would choose to receive a message reminding them to tackle their goals on March 20, 2014 at a higher rate when it was described as a meaningful temporal landmark (the first day of spring) than when it was described as an ordinary day (the third Thursday in March). Indeed, participants in the *meaningful landmark condition* were significantly more likely to choose to receive a reminder to pursue their goal on March 20, 2014 than were participants in the *control condition* (25.61% vs. 7.23%, $\chi^2(1) = 10.18$, $p = .001$; Figure 2). This is a remarkably large response (a 354% increase in sign-ups on March 20th) to a very subtle intervention: relabeling a date many would have already recognized as the first day of spring. Notably, the last day on the list (i.e., March 24, 2014) was selected significantly more frequently than any other days in both conditions (all p-values < .002). The likelihood that the focal date (March 20, 2014) was selected in *the control* condition was not significantly lower than the selection rates of any other days except the last day; thus, the finding that March 20 was selected more frequently in the *meaningful landmark condition* than in the *control*

condition cannot be simply explained by a dislike for March 20ᵗʰ when it was labeled an ordinary day, ("The Third Thursday of March 2014.")

Study 2B

Method. Study 2B extends Study 2A by leveraging a different landmark that is meaningful to a different subject population. In December of 2013, we invited individuals (primarily students) who were signed up to participate in studies at the behavioral lab of a large university in the northeastern United States to take part in an online survey in exchange for a chance to win a $50 Amazon gift card. Study 2B followed a similar procedure to Study 2A with a few key differences. First, participants were asked to describe a goal that they planned to begin pursuing in the *summer* of 2014 (rather than in April 2014) and were offered the opportunity to receive an email reminder sometime in the spring. Second, participants who signed up to receive a future reminder were asked to choose from a list of 14 consecutive dates ranging from May 3, 2014 (Saturday) to May 16, 2014 (Friday). We manipulated our description of May 14, 2014 to read as either "(Wednesday; First Day of [University name]'s Summer Break)" (*meaningful landmark condition*) or "(Wednesday; [University name]'s Administrative Day)" (*control condition*).

We aimed for a sample of 100 participants per experimental condition and invited the entire population of our university's research pool to participate in the hopes of reaching our target sample size. In total, 582 participants responded to our survey. Our sample consisted of 278 participants (62% female, three unspecified; M_{age} = 23, three unspecified) who signed up to receive a reminder and completed the survey. Among

those participants, three participants completed our survey twice, and only their first response was included in our analysis reported below. Our results remain virtually identical if we instead exclude these participants entirely or include each of their multiple responses.

Results. Using a non-overlapping group of participants from the same population ($N = 45$), we confirmed that May 14, 2014 felt more meaningful when it was described to students as the first day of summer break ($M = 5.09$, $SD = 2.45$) than when it was described as their university's "Administrative Day" ($M = 2.00$, $SD = 1.48$, $t(43) = 5.09$, $p < .0001$). Consistent with Study 2A, participants in the *meaningful landmark condition* were significantly more likely to choose to receive a reminder to pursue their goal on May 14, 2014 than were participants in the *control condition* (28.57% vs. 4.35%, $\chi^2(1) = 29.53$, $p < .0001$; Figure 2). Again, we see a remarkably large effect (a 657% increase in sign-ups on May 14th) of a subtle relabeling intervention.

Discussion

By framing otherwise ordinary days as meaningful landmarks that demarcate adjacent seasons (Study 2A) or academic periods (Study 2B), we provide causal evidence that engaging in goal-directed behaviors is more appealing on meaningful landmarks. In an additional, similar study that offered participants an opportunity to receive a future goal reminder (details in Appendix B), we described the same date (October 5) on a list of options as either "the first day after Yom Kippur" (*meaningful landmark*) or "the 278th day of year" (*control*). Consistent with our theory that meaningful landmarks spur goal pursuit, we observed that Jewish participants ($n=86$) rated October 5 as significantly

more meaningful and chose it 25.88% more frequently[21] when it was labeled as the day

after Yom Kippur; our framing manipulation had the opposite effect on non-Jewish

participants (*n*=892). The interaction between our manipulation and whether or not a

participant was Jewish in predicting both the perceived meaningfulness of and choice to

receive a reminder on October 5 was significant (p<.001 and p=.02, respectively).

Study 3

In Study 3, we manipulate the perceived meaningfulness of an otherwise identical

temporal landmark with a guided writing task and then examine participants' engagement

in goal-directed activities.

Method

In early 2014, participants were recruited from Amazon's Mechanical Turk to

take part in a survey about goal pursuit. They were asked first to describe one goal they

had failed to achieve in 2013 and then to indicate what category in a dropdown menu best

captured their goal. Next, participants indicated whether they planned to pursue the

aforementioned goal again this year (i.e., in 2014). Participants who did not plan to

pursue the goal in 2014 exited our survey and were not included in our sample.

Participants who planned to pursue their goal again in 2014 comprised our study sample

and went on to engage in a directed writing task. We aimed for a sample of 100

participants per experimental condition, and we assumed that most but not all participants

would plan to pursue their personal goal again this year. Thus, we invited 250

---

[21] We were unable to recruit our target number of Jewish participants (n=160) before the holiday. Our manipulation directionally increased reminder take-up right after Yom Kippur among Jewish participants but our sample was too small for this effect to reach statistical significance.

participants to participate in our survey. Data collection ceased when total sample size reached our target number. Our final sample included 216 participants (64% female; M_{age} = 29) who planned to pursue their goal again in 2014 and completed our survey.

In the directed writing task, we told participants that we were interested in learning how different people view the start of a New Year, and we randomly assigned them to one of two conditions. In the *meaningful landmark condition*, participants were told that many people view the start of each New Year as a very meaningful day and were then asked to describe three to five reasons why the start of this New Year felt meaningful to them. Participants in the *control condition* were told that many people view New Year's as no different from any other day and were instructed to list three to five reasons why this New Year felt ordinary to them.

Next, we presented participants with information about and links to six different websites that could help them achieve their personal goals, including (a) a website that would allow them to put money on the line that they would forfeit if they failed to follow through on their goal, (b) four popular goal-tracking websites, and (c) a *New York Times* article summarizing insights from recent behavioral science research about how people could increase their chances of achieving their goals. We tracked the number of websites participants clicked on (min = 0, max = 6) as well as the amount of time participants spent reviewing the descriptions of different goal-related websites we provided. The time participants spent reviewing these website descriptions was not successfully recorded for three participants. We included these participants in our other analyses; removing them

does not meaningfully alter our results. At the end of the study, as a manipulation check, participants reported the extent to which New Year's Day 2014 felt meaningful to them.

Results

Our manipulation was effective: New Year's Day 2014 felt more meaningful to participants in the *meaningful landmark condition* ($M = 4.85$, $SD = 1.72$) than to participants in the *control condition* ($M = 3.45$, $SD = 1.65$, $t(214) = 6.09$, $p < .0001$). We predicted that people would engage more in activities designed to facilitate goal pursuit when a past temporal landmark (in this case, New Year's Day) was framed as a more meaningful landmark. First, while most participants did not click on a website (82% clicked zero links in the *meaningful landmark condition* versus 90% in the *control condition*), probably due to an eagerness to complete the task quickly and earn their pay, participants in the *meaningful landmark condition* clicked on three times as many goal-related websites ($M = 0.62$, $SD = 1.54$) as participants in the *control condition* ($M = 0.21$, $SD = 0.75$, $t(214) = 2.4$, $p = .01$), consistent with our prediction. Similarly, participants in the *meaningful landmark condition* spent 46% more time reading our descriptions of these websites ($M = 41.37$ seconds, $SD = 60.09$) than participants the *control condition* ($M = 28.39$ seconds, $SD = 29.62$, $t(211) = 2.00$, $p < .05$).

Discussion

Study 3 demonstrates that having people reflect on the significance of a past temporal landmark increases their engagement in aspirational activities. The next study examines one mechanism that we hypothesize contributes to the motivating effect of meaningful landmarks.

Study 4

Study 4 tests our hypothesis that people feel more psychologically separated from their past, imperfect selves following more meaningful temporal landmarks, which motivates goal pursuit. In this study, participants predicted their own motivation to pursue real, personal goals following a move to a new apartment. We manipulated the meaningfulness of an otherwise identical move by framing it as a first experience.

Method. We recruited 300 participants (41% female, one unspecified; $M_{age} = 32$) online through Amazon's Mechanical Turk. We aimed for a sample of 150 participants per experimental condition. Data collection ceased when total sample size reached our target number. Participants were first asked to think of and briefly describe one goal that they had not succeeded in achieving and would like to pursue in the future. Participants were then asked to imagine that they had just moved to a new apartment that had a similar layout, rent, and commute to work as their previous apartment. They were randomly assigned to imagine either that they had moved for the first time since coming to this city nine years ago (the *meaningful landmark condition*) or that they had moved every year since coming to this city (the *control condition*). Participants then rated how motivated they would be after moving to this apartment to begin pursuing the personal goal that they had described earlier (1 = not at all motivated to 7 = very motivated).

Next, participants were prompted to think about the comparison between their current and past selves. Specifically, they were told the following:

"Most people agree that they have not behaved perfectly in the past (or that their past self has imperfections). There are always some aspects of ourselves and our

lives that we would like to improve. Sometimes our imperfect, past self feels very far away, while at other times our past imperfections feel very close."

They were then asked to rate the psychological distance between their present and imperfect, past selves on three different scales. First, participants were presented with six pairs of Euler circles, which varied in their degree of overlap. Within each pair, one circle represented their imperfect past self one year ago (prior to the apartment move), and the other represented their current self today. Participants were instructed to select whichever pair of circles best reflected their opinion about how far away they would feel today from their imperfect past self, where no overlap between circles meant "extremely far away" and complete overlap meant "extremely close" (adapted from Bartels & Rips, 2010). A second question (adapted from Wilson & Ross, 2001) asked participants to predict the extent to which they would feel *distant* from their imperfect past self (one year prior) (1 = extremely close to 7 = extremely far away). A final question (adapted from Bartels & Rips, 2010) measured participants' perceptions of the extent to which they would feel *different* from their imperfect past self (one year prior) (1 = exactly the same to 7 = completely different). Finally, participants responded to a manipulation check question assessing the perceived meaningfulness of the relocation described.

Results. Again, our manipulation was effective: Participants considered moving to a different apartment to be more meaningful when it was described as their first relocation in nine years ($M = 4.99$, $SD = 1.55$) rather than as their ninth relocation over the same time period ($M = 4.01$, $SD = 1.80$, $t(298) = 5.06$, $p < .0001$). Importantly, participants reported that they would be more motivated to start tackling their personal

goal after a more meaningful (first) relocation ($M = 5.05$, $SD = 1.73$) than after a less

meaningful (ninth) relocation ($M = 4.42$, $SD = 1.89$, $t(298) = 2.98$, $p < .01$). We next

standardized each of our three *psychological disassociation* ratings and averaged them

(with the first rating reverse-coded) to create an index of psychological dissociation

(Cronbach's $\alpha = .88$) in order to conduct mediation analyses. We followed standard

procedures to test whether *psychological dissociation* mediated the relationship between

meaningful landmarks and goal motivation (Preacher & Hayes, 2008). Consistent with

our hypothesized mechanism, participants expected to feel more disconnected from their

imperfect, past self in the *meaningful landmark condition* ($M = 0.15$, $SD = 0.79$) than in

the *control condition* ($M = - 0.15$, $SD = 0.96$, $t(298) = 3.00$, $p < .01$). The composite

psychological disassociation score was a significant, positive predictor ($\beta = 0.24$, $p = .04$)

when we included this measure and a *meaningful landmark condition* indicator variable

in an ordinary least squares regression to predict participants' motivation to pursue their

goals after an apartment move. A 5,000-sample bootstrap analysis showed that the 95%

biased-corrected confidence interval for the size of the indirect effect ($b = 0.07$, $SE =$

0.05) excluded zero [0.003, 0.22]), indicating a significant, positive indirect effect of the

meaningful landmark condition (relative to the *control condition*) through the

psychological disassociation measure.[22]

---

[22] We replicated this mediation result in another scenario study where we asked participants ($n=200$) to forecast a Chinese man's motivation to quit smoking and his perceived separation from his past imperfections following his 36th birthday (see Appendix A). We manipulated this birthday's meaningfulness by either (a) telling participants that it represented the start of a new, 12 year Chinese zodiac cycle (meaningful landmark condition), or (b) withholding this fact (control condition), and we replicated the findings descried above. When this birthday was framed as more meaningful, goal-pursuit was rated as more likely ($p<.0001$), and this effect was mediated by psychological dissociation from past imperfections (95% biased-corrected CI=[0.01, 0.29]).

Discussion

Study 4 shows that the positive effect of a meaningful temporal landmark on a person's motivation to pursue her goals is mediated by an increase in the subjective distance separating her current self from her imperfect past self.

General Discussion

The motivation we need to pursue our aspirations is not a universal constant – sometimes it surges, stimulating goal pursuit, and at other times, it fails us entirely. The present research addresses several major gaps in our understanding of *when* people feel particularly motivated to pursue their goals. First, we demonstrate that meaningful temporal landmarks *cause* people to (a) engage in more goal-directed activities and (b) forecast that their and others' motivation to tackle goals will be heightened. These results hold for landmarks on shared social timetables (e.g., the first day of spring) and in people's life histories (e.g., a relocation); landmarks that have passed (e.g., a recent New Year's) and those yet to come (e.g., the start of next summer's school break). It is well-documented that people are more motivated to pursue the causes that they find personally meaningful (e.g., Grant, 2012; Cryder et al., 2013). However, we demonstrate that imbuing a day (e.g., a date in March, a relocation) with greater meaning can inspire people to pursue an *unrelated* goal (e.g., quitting smoking). Further, we proposed and show that the motivating effect of meaningful temporal landmarks arises in part because more meaningful landmarks create a greater psychological disconnect between a person's current self and her past, inferior self, which motivates goal pursuit. Thus, we elucidate

how landmark-induced discontinuities in individuals' perceptions of time can affect their engagement in goal pursuit.

Our investigation suggests several opportunities for future research. First, future research should explore additional psychological processes that may contribute to the effect documented in this paper, which is likely multiply-determined. For example, it is worth exploring whether meaningful landmarks (e.g., the 40th birthday, a graduation) may disrupt individuals' attention to day-to-day minutiae (Smith, 2014) and stimulate big picture thinking (Liu, 2008; Alter & Hershfield, 2014), which past research has shown promotes a focus on long-term, primary goals (Trope & Liberman, 2003). Another fruitful direction for future research would be to explore the boundary conditions of our findings: disconnecting from one's past self following landmarks may not always be beneficial. For example, people who achieved challenging personal goals in the past may feel more motivated when they remain psychologically connected to those successes.

Our findings also have a number of practical implications. First, for individuals who hope to curtail bad behaviors but struggle with initiating goal pursuit, meaningful temporal landmarks may prevent them from succumbing to "what the hell" rationalizations and falling into vicious cycles of impulsive behavior, thus helping them begin pursuing their aspirations in earnest (Cochran & Tesser, 1996). For managers and policymakers, tools and interventions designed to facilitate the pursuit of long-term goals (e.g., social comparison, commitment devices; Chapman et al., 2014; Schwartz et al., 2014) may be better-received if they are provided following meaningful landmarks. Further, our work suggests that framing an otherwise ordinary day as a temporal

landmark or highlighting the meaningfulness of a temporal landmark may be effective

new "nudges" (Thaler & Sunstein, 2008), capable of bolstering people's interest in

engaging in goal-directed behaviors.

References

Alter, A. L., & Hershfield, H. E. (2014). People search for meaning when they approach a new decade in chronological age. *Proceedings of the National Academy of Sciences of the United States of America, 111*(48), 17066-17070. doi:10.1073/pnas.1415086111

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review, 84*, 191–215

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman

Bartels, D. M., & Rips, L. J. (2010). Psychological connectedness and intertemporal choice. *Journal of experimental psychology. General, 139*(1), 49-69. doi: 10.1037/a0018062

Baumeister, R. F., Vohs, K. D., Aaker, J. L., & Garbinsky, E. N. (2013). Some key differences between a happy life and a meaningful life. *The Journal of Positive Psychology, 8*(6), 505-516

Chapman, G.B., Colby, H.A., Convery, K., & Coups, E.J. (2014). Goals and social comparisons promote walking behavior

Cialdini, R. B. (2007). *Influence: The psychology of persuasion*. New York, NY: HarperCollins

Cochran, W., & Tesser, A. (1996). The "what the hell" effect: Some effects of goal proximity and goal framing on performance. In L. L. Martin & A. Tesser (Eds.), *Striving and feeling: Interactions among goals, affect, and self- regulation*. Hillsdale, NJ: Psychology Press

Colby, H., & Chapman, G. B. (2013). Savings, subgoals, and reference points. *Judgment and Decision Making, 8*(1), 16-24

Cryder, C. E., Loewenstein, G., & Scheines, R. (2013). The donor is in the details. *Organizational Behavior and Human Decision Processes*, *120*(1), 15-23

Dai, H., Milkman, K. L., & Riis, J. (2014). The fresh start effect: Temporal landmarks motivate aspirational behavior. *Management Science, 10*(60), 2563–2582

Festinger, L. (1962). *A theory of cognitive dissonance* (Vol. 2). Stanford, CA: Stanford university press

Grant, A. M. (2012). Leading with meaning: Beneficiary contact, prosocial impact, and the performance effects of transformational leadership. *Academy of Management Journal, 55*(2), 458-476. doi: 10.5465/amj.2010.0588

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology*, *38*(1), 79-109

Libby, L. K., & Eibach, R. P. (2002). Looking back in time: Self-concept change affects visual perspective in autobiographical memory. *Journal of Personality and Social Psychology, 82*(2), 167-179. doi: 10.1037//0022-3514.82.2.167

Liu, W. (2008). Focusing on desirability: The effect of decision interruption and suspension on preferences. *Journal of Consumer Research*, *35*(4). 640-652

Peetz, J., & Wilson, A. E. (2013). The post-birthday world: Consequences of temporal landmarks for temporal self-appraisal and motivation. *Journal of Personality and Social Psychology, 104*(2), 249-266. doi: 10.1037/A0030477

Peetz, J., & Wilson, A. E. (2014). Marking time: Selective use of temporal landmarks as barriers between current and future selves. *Personality and Social Psychology Bulletin, 40*(1), 44-56. doi: 10.1177/0146167213501559

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavioral Research Methods, 40*(3), 879-891. doi:10.3758/BRM.40.3.879

Robinson, J. A. (1986). Temporal reference systems and autobiographical memory. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 159-188). Cambridge, England: Cambridge University Press

Schwartz, J., Mochon, D., Wyper, L., Maroba, J., Patel, D., & Ariely, D. (2014). Healthier by precommitment. *Psychological science*, 0956797613510950

Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory processes. *Psychological Bulletin, 124*(3), 423-442. doi: 10.1037//0033-2909.124.3.423

Smith, I. H. (2014). *The disruption model of inspiration: Toward a general model of "being inspired to act"* (Unpublished doctoral dissertation). University of Utah, Salt Lake City, UT

Soman, D., & Cheema, A. (2004). When goals are counterproductive: The effects of violation of a behavioral goal on subsequent performance. *Journal of Consumer Research*, *31*(1), 52-62

Ross, M., & Wilson, A. E. (2000). Constructing and appraising past selves. In D. L.

    Schacter & E. Scarry (Eds.), *Memory, brain, and belief* (pp. 231-258).

    Cambridge, MA: Harvard University Press

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge*. New Haven, CT: Yale University Press

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review, 110*(3).

    403-421

Wilson, A. E., & Ross, M. (2001). From chump to champ: People's appraisals of their

    earlier and present selves. *Journal of Personality and Social Psychology, 80*(4),

    572-584. doi: 10.1037//0022-3514.80.4.572

Wilson, A. E., & Ross, M. (2003). The identity function of autobiographical memory:

    Time is on our side. *Memory, 11*(2), 137-149. doi: 10.1080/09658210244000324

Wood, W., Tam, L., & Witt, M. G. (2005). Changing circumstances, disrupting habits.

    Journal of Personality and Social Psychology, 88(6), 918-933. doi:

    10.1037/0022-3514.88.6.918

Figure 1. The Meaningfulness Associated with a Given Date Positively Correlates with Motivation to Start a Diet on That Date (Study 1)
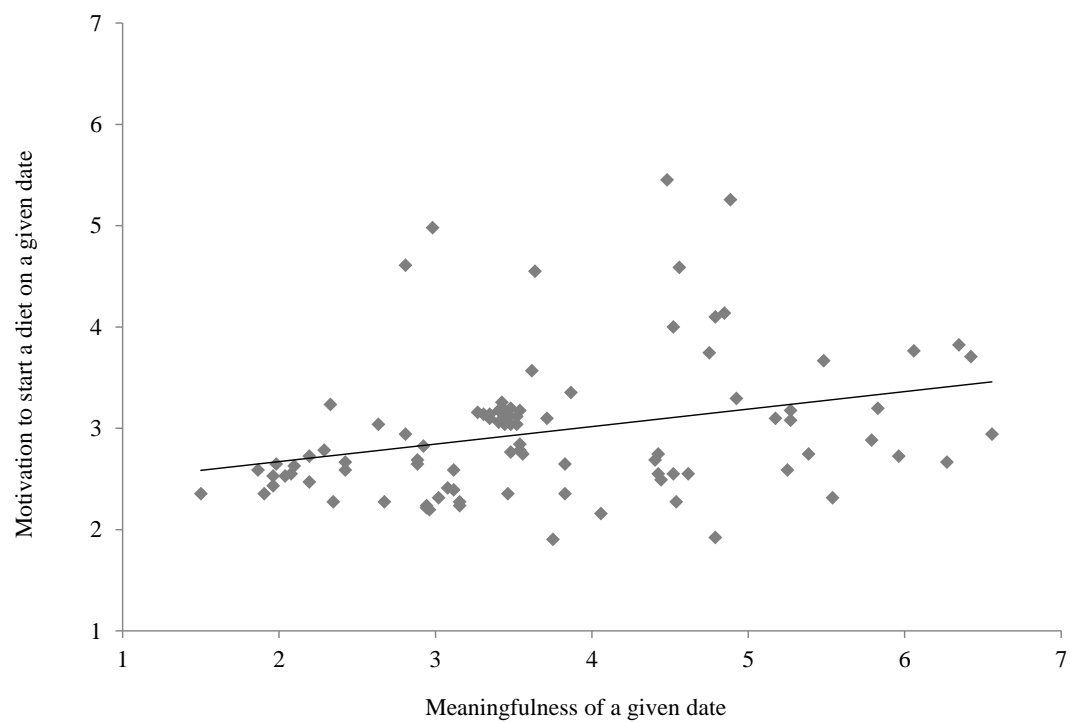
Figure 2. Framing an Otherwise Identical Date as a Meaningful Landmark Increases the Likelihood That Participants Choose to Receive a Reminder About a Goal on the Date in Question (Study 2)
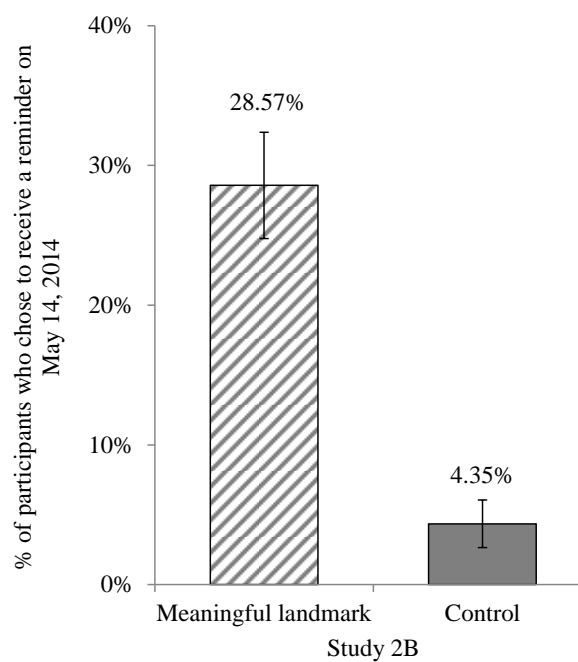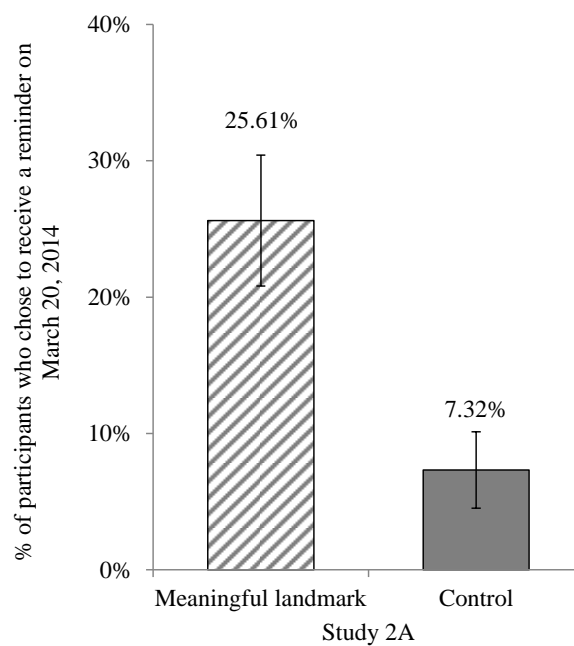
Appendix A: An Additional Reminder Study

In addition to Studies 2A/2B, we conducted another experiment to test whether people are more likely to choose to receive a goal reminder on a day when it is framed as a landmark that they find personally meaningful.

Method

In September of 2014, we recruited participants from three sources: (a) individuals who were signed up to participate in studies at the behavioral lab of a large university in the northeastern United States, (b) individuals who were approached on the campus of the same university, and (c) participants on Amazon's Mechanical Turk. This study followed a similar procedure to Studies 2A/2B with a few key differences. First, participants were asked to describe a goal that they planned to begin pursuing in October or November of 2014. Second, participants who signed up to receive an email reminder were asked to choose from a list of seven consecutive dates ranging from October 1, 2014 (Wednesday) to October 7, 2014 (Tuesday). We manipulated our description of October 5 to read as either "(Sunday; The First Day after Yom Kippur)" (*meaningful landmark condition*) or "(Sunday; The 278th Day of the Year)" (*control condition*). Further, participants were asked to self-report their religious affiliation by selecting one from a list of options ("Christian", "Jewish", "Muslim", "Other(s)", "Unaffiliated", and "Would rather not to answer"). They were also asked to rate the extent to which (a) Yom Kippur and (b) October 5, 2014 felt meaningful to them.

Our sample consisted of 978 participants (62% female; $M_{age} = 27$) who signed up to receive a reminder and completed the survey. Among those participants, 86 individuals

reported that they were Jewish, which was considerably smaller than our target number of Jewish participants (n=160). Our results did not differ significantly by subject population so we collapsed responses across our three subject pools.

Results

First, we confirmed that Jewish participants considered Yom Kippur to be more meaningful (M=5.52, SD=1.65) than non-Jewish participants (M=1.59, SD=1.14, $t(975)$=29.22, p<.0001). Then we confirmed that October 5, 2014 felt *more* meaningful to Jewish participants when it was described as the first day after Yom Kippur (M=4.84, SD=2.02) than when it was described as the 278th day of the year (M=2.71, SD=2.02, $t(84)$=4.69, p<.0001), whereas we observed the opposite pattern among non-Jewish participants ($M_{meaningful\_landmark\_condition}$=1.89, SD = 1.57 vs. $M_{control\_condition}$=2.48, SD=1.84, $t(890)$=5.14, p<.0001). The interaction between our manipulation and whether or not a participant was Jewish in predicting the perceived meaningfulness of October 5 was significant ($p < .001$).

When choosing when to receive a reminder, Jewish participants in the *meaningful landmark condition* were (directionally) more likely to choose October 5 than participants in the *control condition* (32.73% vs. 25.81%, $\chi^2(1)$=0.45, p=.50); however, non-Jewish participants in the *meaningful landmark condition* chose October 5 less frequently than did participants in the *control condition* (8.51% vs. 19.26%, $\chi^2(1)$=21.37, p<.0001). The interaction between our manipulation and whether or not a participant was Jewish in predicting an individual's choice to receive a reminder on October 5 was significant (p=.02).

Appendix B. An Additional Study that Tests the Proposed Underlying Mechanism

In addition to Study 4, we conducted another study to test our hypothesis that people feel psychologically more separated from their past, imperfect selves following more (versus less) meaningful temporal landmarks, which motivates goal pursuit. In this study, participants predicted a hypothetical man's motivation to pursue a challenging goal following a birthday. We manipulated the meaningfulness of an otherwise identical birthday by imbuing it with cultural relevance.

Method

We recruited 200 Amazon's Mechanical Turk participants (44% female, $M_{age}=$ 32). Participants were asked to imagine that a man named Chang who lived in China had just celebrated his 36th birthday the day before. They were told that Chang had wanted to quit smoking for a long time but had never succeeded. We randomly assigned participants to one of two experimental conditions. In the *meaningful landmark condition*, they were introduced to the concept of the 12-year Chinese zodiac cycle and were told that Chang's 36th birthday represented the beginning of his fourth zodiac cycle. In the *control condition*, participants were not told about the Chinese zodiac cycle. Participants in both conditions rated how motivated Chang would be to quit smoking after celebrating his 36th birthday (1=not at all motivated to 7=very motivated). Next, participants used the same scales described in Study 4 to rate the psychological distance between Chang's present self and his imperfect, past self, imagined two years prior to his 36th birthday. Finally, participants evaluated the meaningfulness associated with Chang's 36th birthday as a manipulation check (1=not at all to 7=very much).

Results

We first confirmed that our manipulation was effective: participants believed

Chang's 36th birthday would be more meaningful to him in the *meaningful landmark*

condition, where his 36th birthday was described as corresponding to the start of a new

zodiac cycle ($M_{meaningful\_landmark\_condition}$=5.68, *SD*=1.22 vs. $M_{control\_condition}$=4.56, *SD*=1.49,

t(198)=5.88, *p*<.0001). Further, we found that participants believed Chang would be

more motivated to quit smoking in the *meaningful landmark condition* (*M*=4.93,

SD=1.31) than in the *control condition* (*M*=4.02, *SD*=1.64, *t*(198)=4.35, *p* <.0001).

As in Study 4, we standardized the three measures of *psychological disassociation*

and averaged them to create an index of psychological dissociation (Cronbach's α = .77).

Participants believed that Chang would feel more dissociated from his imperfect, past self

in the *meaningful landmark condition* (*M* = 0.18, *SD* = 0.82) than in the *control condition*

(*M* = -0.18, *SD* = 0.80, *t*(198) = 3.15, *p* < .01). The composite *psychological*

disassociation score was a significant, positive predictor (β=0.32, *p*=.01) when we

included this measure and a *meaningful landmark condition* indicator variable in an

ordinary least squares regression to predict the extent to which participants predicted

Chang would be motivated to quit smoking after his 36th birthday. A 5,000-sample

bootstrap analysis showed that the *psychological disassociation* score mediated the

relationship between the *meaningful landmark condition* indicator variable and

motivation (*b*=0.12, *SE*=0.07, 95% biased-corrected CI=[0.01, 0.29]).

CHAPTER 3

A DOUBLE-EDGED SWORD: HOW AND WHY

RESETTING PERFORMANCE METRICS AFFECTS FUTURE PERFORMANCE

Hengchen Dai

ABSTRACT

In organizations, metrics that track individuals' performance are often reset to zero at the start of a new performance period (e.g., a new month for salespeople, a new season for sports players). Additionally, organizations may disregard individuals' historic performance records due to various events (e.g., a relocation to a different branch office, the implementation of a new incentive program that only rewards high future performance). I examine the impact on individuals' future performance of tracking that performance without incorporating records of their past performance (a phenomenon I refer to as a *performance reset*). I propose that when individuals believe their past performance was poor, resets will improve their performance by boosting self-efficacy and commitment. However, I expect that when individuals believe their past performance was strong, performance resets will hurt their performance by decreasing commitment without increasing self-efficacy. One archival field study involving professional baseball players and four laboratory experiments support this hypothesized relationship between performance resets, past performance, and future performance; these studies also provide preliminary evidence that self-efficacy mediates this relationship.

Keywords: Performance Resets, Motivation, Self-efficacy, Commitment

Introduction

To motivate employees, organizations carefully track their performance and allocate resources (e.g., rewards, time, support) accordingly. Metrics that track individuals' performance are often reset to zero at the start of a new performance period. For example, a salesperson's tally of cumulative sales may be reset to zero at the beginning of each month, and many performance statistics for professional sports players are reset each season. In addition to these routinized, periodic resets of performance metrics, performance records may be temporarily or permanently wiped clean as a side effect of numerous organizational events. For instance, after being traded to a different league in the middle of a regular season, a Major League Baseball player's season-to-date statistics will be reset to zero at the start of his first game on his new team. As a more extreme example of disregarding past performance, a manager who launches a program designed to reward the employee with the highest attendance in the following quarter is implicitly treating past attendance records as irrelevant to employees' future performance appraisals. These and other similar examples raise a critical question about the impact of resetting performance records: When individuals' performance is tracked from a new starting point without involving their past performance, can such a *performance reset* affect their subsequent performance, and if so, how and why?

Existing research that examines how past performance affects future performance via factors such as self-efficacy and goal commitment (e.g., Locke, Frederick, Lee, & Bobko, 1984; Bandura, 1997; Vancouver, 1997; Fishbach & Dhar, 2005; Heggestad & Kanfer, 2005) has looked at past performance in a somewhat static way. Specifically, our

past performance is viewed as something that is core to our being, signaling what we value and reflecting what we can achieve in the future. Such a perspective overlooks organizational practices that track individuals' performance from a new starting point and thus impose a separation between past and future performance. However, feeling psychologically distant from past successes or failures may have important implications for individuals' motivation and subsequent performance. This paper examines the conditions under which having a clean slate on performance records may spur versus hamper individuals' motivation to perform well. Addressing this issue will also advance our understanding of organizational factors that may encourage people to improve their future performance in some situations but decrease their motivation in others.

Existing research on motivation and performance management has established that self-efficacy and commitment to performing well significantly and positively affect future performance (e.g., Bandura, 1997; Klein & Kim, 1998; Locke, Latham, & Erez, 1988; Tolli & Schmidt, 2008). Drawing on the literatures on performance feedback, self-efficacy, commitment, and motivation, I propose that the effects of a performance reset on future performance should depend on how an individual views her past performance. Further, I propose that the relationship between these variables can be explained by changes in individuals' self-efficacy and commitment. Specifically, I propose that when individuals believe their past performance was poor, resetting their performance records is likely to boost their self-efficacy because they now view past failures as less indicative of their future ability. In addition, a clean slate induced by a performance reset may increase their commitment to performing well because they now have a new opportunity

to establish their capability. However, when individuals believe their past performance was strong, performance resets are unlikely to improve their self-efficacy because they have already formed a positive view of their capability of excelling at work. Rather, performance resets may decrease their commitment because once past achievements are set in stone, the value of proving themselves is reduced. In light of the positive effects of self-efficacy and commitment on performance, I predict that performance resets will improve future performance when individuals have weak past performance but dampen future performance when they have strong past performance. Altogether, this paper provides a theoretical framework that explicates how and why resets affect future performance.

I test my hypotheses in five studies. Study 1 examines the effects of performance resets in the context of Major League Baseball (MLB). I find that a performance reset improves the rate at which batters successfully hit when at bat if they performed poorly prior to the reset but harms hitting performance among batters who performed relatively well prior to the reset. In Studies 2-4, I provide *causal* evidence that performance resets improve performance when people view their past performance as poor but decrease performance when people think positively of their past performance. These studies experimentally manipulate the occurrence of performance resets. Study 4 further demonstrates that self-efficacy significantly mediates the relationship between performance resets, past performance, and future performance. Study 5 provides additional support for my hypotheses by showing that people rationally differentiate the value of performance resets based on their perceptions of their past performance.

Specifically, individuals who view their past performance less positively are more likely to elect to reset their performance statistics than individuals who view their past performance more positively. In all of my studies, performance resets are compared with a control condition where past performance is carried over the current performance period. Thus, my findings that performance resets improve future performance for individuals with weak past performance but harm future performance for individuals with strong past performance cannot simply be explained by regression to mean.

Conceptual Framework

Performance Resets Create a Separation from Past Outcomes That Influences Future Performance

The perceived continuity between past, current, and future selves, states, and outcomes is malleable (e.g., Peetz & Wilson, 2013; Bartels & Urminsky, 2011). Simple environmental cues may lead people to perceive their past and future states as separate (Peetz & Wilson, 2013) and evaluate past and future outcomes independently (Read, Loewenstein, & Rabin, 1999). For example, prior financial losses (e.g., losses at the gambling table) have a weakened influence on people's subsequent risk-taking behavior when the losses are realized (e.g., cashing out after a gambling loss). This is because realizing prior losses closes the associated mental accounts and resets the reference point against which people compare subsequent prospects (Imas, 2014). Relatedly, landmark events that demarcate the beginning of a new period on calendars (e.g., the beginning of the year, the Spring equinox) or in personal histories (e.g., birthdays, relocations) lead people to feel psychologically distant from their past and future selves (Dai, Milkman,

Riis, 2015; Peetz & Wilson, 2013). This stream of literature suggests that salient transitions between two adjacent periods create boundaries between mental categories that lead elements in one period to be perceived as dissimilar to elements in the other period. Thus, we expect that resetting individuals' performance metrics to zero creates a mental separation from their past performance and propels them to appraise their future outcomes relative to a clean slate.

How will performance resets affect individuals' subsequent performance? What determines whether individuals respond to performance resets by either redoubling or relaxing their effort? Recent research on the "fresh start effect" (Dai, Milkman, & Riis, 2014, 2015) demonstrates that temporal landmarks—including personally meaningful life events (e.g., birthdays) and time dividers on calendars (e.g., holidays, Mondays)— motivate people to tackle their self-improvement goals. One reason is that people feel distant from their past failures following temporal landmarks and thus feel more confident about their current selves (Dai et al., 2015). Extending this "fresh start effect" to the context of performance resets suggests that when individuals do not feel they have performed well, having a clean slate on their performance records may motivate them to work harder. However, what if people feel that they have performed well? I argue that they may feel less motivated to exert effort if their performance is tracked from a new reference point because they have already proved their ability to excel at the task and may find the prospect of starting a new performance record unappealing. That is, I propose that the effects of performance resets on future performance should depend on how people view their past performance. Specifically, I hypothesize the following:

Hypothesis 1: Past performance will moderate the effects of performance resets on future performance. Specifically, resets will increase future performance when individuals view their past performance as poor but decrease future performance when individuals view their past performance as strong.

My proposition is in line with the results of Gubler, Larkin, and Pierce (2014). They studied an awards program that offered employees with perfect attendance in a given month an opportunity to receive a prize; the award program thus implicitly disregarded employees' historic attendance records. Comparing the factory that implemented this award program with similar factories that did not, Gubler et al. (2014) found that employees with the highest tardy rates prior to the introduction of the award program decreased their tardy rates after the program was implemented. They also found that employees with moderate pre-intervention tardy rates did not decrease their tardy rates and employees who were never tardy in the previous 12 months increased their likelihood of taking planned absences and lowered their productivity by 6 percent.[23] These results provide suggestive evidence that past performance may alter the effects of performance resets on future performance. However, in Gubler et al. (2014), the award program also emphasized a performance metric that was previously overlooked by employees (i.e., attendance) and intensified competition among coworkers. In the current research, I examine the impact of performance resets without introducing new performance metrics or altering competition.

---

[23] Gubler et al. (2014) speculated that employees with perfect pre-intervention attendance increased their absence rates and decreased their productivity because they felt it unfair that their past attendance was disregarded and that those with poor pre-intervention attendance had the same chance to win the award.

In the following sections, I theorize that two significant determinants of performance—individuals' task-specific self-efficacy and their commitment to performing well—drive the psychological processes through which past performance moderates the effects of performance resets on future performance.

Performance Resets and Past Performance Influence Self-efficacy

Self-efficacy is a person's belief in his ability to organize and execute the actions required to succeed at a given task (Bandura, 1997). Self-efficacy has been shown to predict actions and future performance. Compared to those with low self-efficacy, individuals with high self-efficacy select more challenging tasks (Bandura & Schunk, 1981), set more challenging goals (Brown, Jones, & Leigh, 2005; Tolli & Schmidt, 2008), exert more effort (Schunk & Hanson, 1985; Schunk, Hanson, & Cox, 1987; Schmidt & DeShon, 2010), work more persistently when faced with difficulties (Multon, Brown, & Lent, 1991), and eventually achieve higher performance (Bandura & Locke, 2003; Brown et al., 2005; Stajkovic & Luthans, 1998; Stefano, Gino, Pisano, & Staats, 2014). One determinant of self-efficacy is an individual's past performance: Past successes indicate that you are capable of harnessing the resources required to perform well and elicit high self-efficacy, whereas past failures cast doubt on future accomplishments and decrease self-efficacy (Bandura, 1991, 1997; Locke et al., 1984).

Notably, the impact of past failures on self-efficacy is weakened when people feel distant from past failures and do not consider them to be reflective of their current and future capabilities. For example, Libby and Eibach (2002) showed that when dieters use a third-person perspective to recall past episodes of overindulgent eating, they become

more optimistic about their ability to refrain from overeating during their next Thanksgiving dinner because they no longer view their undesirable past actions as a reflection of their new, current selves. Similarly, attributing prior failures to externally-controlled factors (as compared to internally controlled factors) has been shown to increase self-efficacy (Donovan & Williams, 2003; Tolli & Schmidt, 2008). More broadly, Polivy and Herman (2002) contend that many individuals engage in frequent attempts at self-improvement despite repeated past failures because they sustain their optimism by attributing failures to external factors and their past selves. Integrating this line of research with the notion that performance resets separate past and future performance, I propose that resets will boost self-efficacy when individuals view their past performance as poor. This is because resets may make those individuals feel untarnished by their unsuccessful past performance and more confident about their ability moving forward.

Among individuals who performed well in the past, we might expect performance resets to decrease self-efficacy if resets psychologically separate them from their successful past selves and make them feel less positively about their current selves. However, I argue that performance resets should have no effect (rather than a detrimental effect) on self-efficacy when individuals have strong past performance. The desire to feel competent is a basic human motivation (White, 1959; Ryan & Deci, 2000), and people are good at claiming ownership of past achievements (e.g., by pulling past achievements closer to their current selves) in an effort to maintain a positive self-view (Ross & Wilson, 2002; Wilson & Ross, 2001). Thus, I expect that even when past achievements

are no longer part of performance appraisals following a reset, individuals will still be able to maintain a high level of self-efficacy due to their memories of past successes.

Combining these arguments, I propose the following hypothesis:

Hypothesis 2: Past performance will moderate the effects of performance resets on self-efficacy. Specifically, the positive relationship between resets and self-efficacy will be stronger when individuals view their past performance as poor than when they view their past performance as strong.

Performance Resets and Past Performance Influence Commitment

In addition to individuals' self-efficacy, their commitment to achieving a high performance level also positively influences the effort they subsequently allocate to the task and their final performance (e.g., Fishbach & Dhar, 2005; Klein & Kim, 1998; Locke et al., 1988; Riedel, Nebeker, & Cooper, 1988; Zhang & Huang, 2010). One significant determinant of individuals' commitment to a performance goal is the general value of the goal (Locke et al., 1988; Zhang & Huang, 2010): The more people value the goal of achieving a performance level, the more committed they are to taking actions and exerting effort.

I propose that the effects of performance resets on individuals' commitment to performing well on a task depend on how they view their past performance. When individuals view their past performance positively, performance resets turn strong past performance into a successful record that is set in stone; by contrast, without performance resets, past performance becomes part of continuing performance appraisals but does not stand alone as proof of prior achievements. Consequently, performance resets may lead

120

those with strong past performance to feel that they have already established their ability to excel at the focal task and that it is less valuable to achieve strong future performance and prove themselves again. Since the perceived value of a goal is positively correlated with individuals' commitment (Locke et al., 1988; Zhang & Huang, 2010), performance resets may reduce commitment among individuals who previously performed well.

However, I expect performance resets to increase commitment among individuals who believe their past performance was weak. Without performance resets, weak past performance is carried over to the current performance period and drags down overall performance, making it difficult for individuals to catch up even if they work very hard in the future. This may in turn reduce their engagement in the task (Cochran & Tesser, 1996). By contrast, a performance reset offers individuals an opportunity to build up their performance record from scratch. Performing well after a reset will allow individuals with weak past performance to establish their capability using their new record. Thus, performance resets may make it more valuable for those individuals to improve their performance and increase their commitment to a task. Consider a baseball player who has been performing poorly in the first two months of a regular season. He may realize that he has little chance to save his overall seasonal performance statistics regardless of how hard he tries in the next few months because his past performance is a burden. As a result, he may not feel committed to improving his performance. However, if he is traded to another league and has his season-to-date statistics reset, he has a new opportunity to prove himself and may feel more committed to achieving strong subsequent performance.

Combining these arguments, I propose the following hypothesis:

Hypothesis 3: Past performance will moderate the effects of performance resets on commitment. Specifically, resets will boost commitment when individuals view their past performance as poor but harm commitment when individuals view their past performance as strong.

Performance Resets and Past Performance Influence Future Performance via Self-efficacy and Commitment

Self-efficacy and commitment have been shown to positively influence future performance. Research has simultaneously examined the effects of self-efficacy and commitment on performance, suggesting that self-efficacy and commitment have independent explanatory power (Sue-Chan & Ong, 2002). Thus, I expect that the psychological processes through which performance resets and past performance affect self-efficacy and commitment will have further implications for individuals' subsequent performance. Specifically, I propose that for individuals with weak past performance, a performance reset should improve future performance by boosting their self-efficacy and commitment to the task in question; however, a performance reset should decrease subsequent performance for individuals with strong past performance by reducing their commitment to the task in question without changing their self-efficacy. I formally state my hypothesis as follows:

Hypothesis 4: Both self-efficacy and commitment will mediate the relationship between performance resets, past performance and future performance. Specifically, past performance will moderate the effects of performance resets on

self-efficacy and commitment (i.e., Hypotheses 2 and 3), which will positively influence future performance.

Overview of Studies

I tested Hypotheses 1-4 across five studies: an archival study based on real performance outcomes gathered from a field setting and four laboratory experiments. In Study 1, I collected performance data on professional baseball players who were traded in the middle of the regular baseball season, taking advantage of the fact that inter-league trades reset season-to-date statistics but within-league trades carry over the performance statistics players achieved before they were traded. I examined how hitting performance — measured by the probability of successfully getting a hit when at bat—changes as a function of (a) whether a trade occurred between or within leagues and (b) a player's hitting performance prior to the trade. In Study 2, I experimentally manipulated the presence of a performance reset and tested whether performance resets have a different impact on individuals' performance on a cognitive task depending on past performance. Studies 3 and 4 used a task that requires motor skills and independently manipulated (a) the presence of a performance reset and (b) individuals' perceptions of their past performance. This allowed me to provide a *causal* test of whether an individual's view of her past performance alters her response to a performance reset. Study 4 further examined whether past performance and performance resets affect future performance by altering self-efficacy and commitment. Study 5 tested my hypothesis in a new way and examined whether people choose to reset their performance statistics at a higher rate if they believe their past performance was poor than if they believe their past performance was strong.

Study 1

In Study 1, I analyze data on 40 years of Major Baseball League games and examine the performance of baseball players who were traded in the middle of the regular season. I take advantage of the fact that players' season-to-date statistics are reset after they are traded to a different league but are carried over after they are traded to a different team in the same league. I investigate whether the effects of being traded across leagues (versus being traded within a league) on players' post-trade performance depend on players' recent performance prior to a trade.

Setting

Major League Baseball (MLB) offers an appropriate setting to test Hypothesis 1 for at least two reasons. First, MLB consists of two leagues, the American League (AL) and the National League (NL).[24] When a player is traded to the other league in the middle of a regular season, his season-to-date statistics will start from zero, and his statistics in the previous league will not be counted toward his totals in the new league; but if a player switches to another team in the same league, his statistics will carry over. Statistics are reset following inter-league trades for league leaderboards and awards purposes because (a) league leaderboards and awards are league specific and (b) it is deemed inappropriate for a player to lead in a league in which he has never or barely played. For both inter-league and within-league trades, players' performance before and after a trade counts toward their career statistics. The second factor that makes MLB an appropriate setting is

---

[24] The two leagues use the same set of rules and regulations, with the exception that the AL operates under the designated hitter rule, and the NL does not. The designated hitter rule allows baseball teams to have one player (known as the designated hitter) who bats in place of the pitcher. Without this rule, the pitcher in a team must bat.

that although baseball is a team sport, each team member performs more or less on his own. Thus, each team member's contributions are relatively independent and can be rather objectively assessed (Mandelbaum, 2005).

Data

Play-by-play data for all Major League Baseball players from 1975 to 2014[25] were obtained based on Retrosheet's (n.d.) regular-season event files. I also collected information about trades that occurred during the same period based on Retrosheet's (n.d.) transaction database. I focused on trades that were arranged between a team's first game and its last game in a regular season (hereafter, *mid-season trades*) because only mid-season inter-league trades result in performance resets. Since each player can appear in multiple seasons in this data, I use the term *player-season* to capture a given player in a given season. I focused on the effects of inter-league trades on batters[26] and identified 1,408 player-seasons of 1,116 unique players who batted in at least one game before and after a mid-season trade.[27]

---

[25] I started my observation period from 1975 primarily because event files on Retrosheet are missing a few games for most seasons prior to 1974 (Retrosheet, n.d.). In addition, free agency, which allows players to openly negotiate their contract with any team after their current contract has expired, became available in the mid-1970s and has increased player mobility.

[26] I focused on batters rather than pitchers for two reasons. First, as explained earlier, pitchers must bat in the NL but do not have to bat in the AL. Thus, inter-league trades change whether or not pitchers must bat in a game. Second, pitchers in the NL can often get better outcomes when pitching to the opposing team's pitcher, who is usually not a very good batter; but pitchers in the AL normally do not get to pitch to the opposing team's pitcher. As a result, pitching statistics in the NL tend to be better than pitching statistics in the AL (Schwarz, 2007). In summary, the AL's adoption of the designated hitter rules may cause inter-league trades and within-league trades to differently change pitchers' statistics for reasons other than the reset of performance statistics. See Appendix A for results based on pitchers' data.

[27] For the 2.45% of MLB players who were traded twice or three times in the middle of a season, I examined their first mid-season trade, thereby avoiding using those players' observations after their first trade as the pre-trade baseline for their subsequent trades in the same season. My results are robust if I exclude the players who were traded more than once in the middle of a season.

For two reasons, my performance measure was whether or not a batter successfully hit the ball each time he was credited with an at bat. First, batting average, which the Baseball Almanac (n.d.) refers to as "easily the most common statistic in baseball and the most understood," is calculated as the number of times a batter successfully hits the ball divided by the number of at bats. Second, a box score (or the summary of a game's statistics) at minimum includes the number of at bats and hits[28] each batter was credited with in a game. Consistent with my measure of performance, I operationalized batters' *past performance* by calculating their batting average during the pre-trade period of the season. To ensure a sufficiently granular and meaningful batting average, I restricted my sample to players who had at least 100 at bats during the pre-trade period, which reduced my sample to 706 player-seasons of 558 unique players. Our results are robust if I relax this exclusion criterion. Across the 706 trades involved in my main analysis, 41.93% were inter-league trades.

Variables

Dependent variable.

Hitting indicator. Each time a batter was credited with an at bat, *hitting indicator* equaled 1 if the batter successfully hit the ball and 0 if the batter did not hit the ball ($M = 0.26$, $SD = 0.44$).

Independent variables.

---

[28] Other batting statistics are not necessarily included in a box score. For example, base on balls (the number of times a batter is entitled to reach first base without the possibility of being put out because he receives four pitches that the umpire calls *balls*) may or may not be included in a box score.

Inter-league indicator. When a player was traded across leagues, *inter-league indicator* equaled 1; otherwise, it equaled 0 ($M = 0.42$, $SD = 0.49$).

Pre-trade batting average. For each batter in a season, I calculated the player's batting average ($M = 0.260$, $SD = 0.032$) from the first until the last game he played for the team from which he was traded. The calculations were validated against data from Baseball-Reference.com, a website that is often used by major media organizations and baseball broadcasters as a source for baseball statistics.

Post-trade indicator. This dummy variable equaled 1 if a given observation occurred after a given player was traded and otherwise equaled 0.

Analysis strategy

I relied on the following ordinary least squares (OLS) regression specification to predict the hitting indicator for at bat *i* for player-season *j*:

$$hitting\ indicator_{ij} = \alpha + \beta(post - trade\ indicator)_{ij} +$$

$$\lambda[(post - trade\ indicator)_{ij} * (inter - league\ indicator)_j] +$$

$$\tau[(post - trade\ indicator)_{ij} * (pre - trade\ batting\ average)_j] +$$

$$\zeta[(post - trade\ indicator)_{ij} * (inter - league\ indicator)_j *$$

$$(pre - trade\ batting\ average)_j] + \delta_j + \varepsilon_{ij},$$

where the two-way and three-way interaction between the post-trade indicator and the other two independent variables (the inter-league indicator and pre-trade batting average) were included. I included player-season fixed effects δ_j, allowing me to estimate how hitting performance changes after a trade while controlling for the player's average performance in a given season. Notably, the inter-league indicator and pre-trade batting

average had a fixed value for a given player-season and were collinear to player-season

fixed effects. Thus, the main effects of inter-league (versus within-league) trades and pre-

trade batting average as well as their interaction were not included in the regression. I

clustered the standard errors at the player-season level.

Results

       Descriptive statistics and correlations are reported in Panel A of Table 1. As

summarized in Panel B in Table 1, inter-league trades and within-league trades were

comparable on relevant observables.

       Summary statistics showing the effects of inter-league trades on post-trade

hitting probability. For ease of exposition, I graphically present the results in Figure 1

by focusing on the bulk of the observations that happened either within 90 days prior to a

trade or within 90 days after a trade.[29] To understand how players' responses to inter-

league versus within-league trades change over time,[30] I separate both the pre-trade

period and the post-period period into three 30-day intervals. Figure 1 depicts the average

probability of hitting at bat as a function of (a) whether a player's pre-trade batting

average was one standard deviation above the mean (i.e., 0.292) or below the mean (i.e.,

0.228), (b) whether the trade in question was an inter-league or within-league trade, and

(c) each of the 30-day intervals prior to versus after the trade. For both players with a

high pre-trade batting average and players with a low pre-trade batting average, the

---

[29] More than 80% of players played the last game in the regular season for their new team within 90 days after they were traded. Similarly, approximately 75% of players played the first game in the regular season for their previous team within 90 days before they were traded.

[30] I did not expect the effects of a performance reset caused by an inter-league trade to persist throughout the season. Instead, I expected that the effects of a performance reset would arise early on and fade over time.

probability of hitting at bat did not significantly differ between inter-league and within-league trades across the three 30-day intervals before trades (all ps > .49). This suggests that batting performance measured as the probability of hitting at bat was comparable between inter-league and within-league trades prior to a trade.

My theory predicts that the lower a batter's pre-trade batting average, the more an inter-league trade would improve his post-trade batting performance over and above a within-league trade. To test this prediction, I focus on the post-trade period. For players whose pre-trade batting average was one standard deviation below the mean, their average probability of hitting was significantly higher during the period of 0-30 days and 30-60 days following inter-league trades ($M_{0-30days} = 0.263$, $SD = 0.440$; $M_{30-60days} = 0.270$, $SD = 0.444$) than following within-league trades ($M_{0-30days} = 0.236$, $SD = 0.425$; $M_{30-60days} = 0.243$, $SD = 0.429$; both ps < .03). The benefits of inter-league trades among those players disappeared after 60 days following a trade.

However, the effects of inter-league trades (relative to within-league trades) turned in the opposite direction among players whose pre-trade batting average was one standard deviation above the mean. Specifically, their average probability of hitting across all of three 30-day intervals was directionally lower following inter-league trades ($M_{0-30days} = 0.287$, $SD = 0.452$; $M_{30-60days} = 0.279$, $SD = 0.449$; $M_{60-90days} = 0.254$, $SD = 0.436$) than following within-league trades ($M_{0-30days} = 0.292$, $SD = 0.455$; $M_{30-60days} = 0.283$, $SD = 0.450$; $M_{60-90days} = 0.254$, $SD = 0.436$; $M_{60-90days} = 0.279$, $SD = 0.449$; all ps > .26).

In summary, Figure 1 shows that inter-league trades improved post-trade hitting performance over and above within-league trades if players' pre-trade batting average was low, but the effects of inter-league trades reversed if players' pre-trade batting average was high. These patterns were prominent during the first 60 days following a trade but faded after 60 days.

Regression analyses showing the effects of inter-league trades on post-trade hitting probability. The analyses reported above were based on raw data without controlling for any factors and did not address the concern that a player-season has repeated observations. Next, I turned to the regression model described in the Analysis Strategy section, controlling for player-season fixed effects and clustering standard errors at the player-season level.

As shown in Model 1 of Table 2, the positive and significant coefficient on the *post-trade indicator* ($p < .001$) means that the probability of hitting at bat on average improved after a within-league trade. Further, hitting probability on average improved more substantially after an inter-league trade than after a within-league trade, as indicated by the positive and significant coefficient on the interaction between the *post-trade indicator* and the *inter-league indicator* ($p = .014$).[31] Importantly, the negative and significant three-way interaction between the *post-trade indicator,* the *inter-league indicator*, and *pre-trade batting average* ($p = .014$) suggests that the lower a player's pre-trade batting average, the more likely inter-league trades were to improve hitting

---

[31] The negative and significant interaction between the *post-trade indicator* and *pre-trade batting average* suggests that the lower a player's pre-trade batting average, the more their hitting probability improved after a trade. This may simply reflect regression to mean and is irrelevant to my hypothesis.

probability above and beyond within-league trades. Specifically, when a player's pre-trade batting average increases by one standard deviation (i.e., 0.032), the additional improvement in hitting probability induced by an inter-league trade would on average decrease by 0.008, which is equivalent to 3% of the average hitting probability of an average player in my dataset. In fact, compared with within-league trades, inter-league trades decrease post-trade hitting performance ($p = .039$) for players whose pre-trade batting average was two standard deviations above the mean (i.e., 0.324). These findings are consistent with the patterns observed in Figure 1 as well as Hypothesis 1.[32]

Placebo test addressing alternative explanations. Inter-league trades and within-league trades may differ in aspects other than the reset of performance statistics. Importantly, the focus of my analysis is to test whether pre-trade batting average affects the effects of inter-league trades on post-trade batting performance rather than to test the main effects of inter-league trades. Thus, other inherent differences between inter-league and within-league trades would only create alternative explanations if these differences cause post-trade hitting probability to change differently for players with a low versus high pre-trade batting average.

To address the possibility that my findings are explained by differences between the two types of trades other than performance resets, I take advantage of the fact that inter-league trades occurring between seasons do not reset performance statistics. I applied the same analysis strategy described above to analyzing trades that occurred

---

[32] In an alternative regression, I replaced the post-trade indicator with four post-trade indicators, each of which represented the 0-30 days, 30-60 days, 60-90 days, and beyond 90 days after a trade, respectively. The regression results are consistent with Figure 1 and Hypothesis 1. See Appendix C for details.

between seasons but did not see the same patterns predicted by my theory (Model 2 in Table 2). This suggests that the different impact of mid-season inter-league trades on players with a low versus high pre-trade batting average are unlikely to be explained by something inherent to such trades besides the introduction of a performance reset. See details about this placebo test in Appendix B.

Robustness checks. I conducted a number of robustness checks to ensure that the findings above are not the spurious results of the regression specifications and sample selection criteria. See Appendix D for detailed explanations and regression results. The findings are robust if:

- I use a logistic regression to predict the hitting indicator;

- I relax or strengthen the criterion I used to ensure that pre-trade batting average is granular and reliable;

- I only include players having more than 100 at bats both before and after a trade;

- I only include players who were not also a pitcher in a given season;

- I only include players who had one mid-season trade in a season; and

- I control for the Batting Park Factor of the team corresponding to each observation.[33]

---

[33] A ballpark's Batting Park Factor captures the extent to which this ballpark is favorable to batters versus pitchers. It is plausible that batters who have a low pre-trade batting average and are traded across leagues are more likely to be traded to a team whose ballpark is friendly to batters and thus seemingly improve their performance, compared with (a) batters with a low pre-trade batting average who are traded within a league and (b) batters with a high pre-trade batting average. If this is true, then pre-trade batting average may moderate the effects of inter-league trades on hitting probability for reasons other than the reset of performance statistics. However, this speculation is not supported by my ancillary analysis because players with a low pre-trade batting average who are traded across leagues are not more likely than others to switch to a team with a more batter-friendly ballpark. See Appendix E for details about this analysis. To further address this alternative account, I controlled for Batting Park Factors in my robustness checks.

Discussion

Study 1 examines a high-stakes field setting where performance resets are naturally induced by mid-season interleague trades. Consistent with Hypothesis 1, I find that trading across leagues (relative to trading within leagues) improves players' probability of hitting at bat among those who had a low batting average prior to a trade; however, the effects of inter-league trades reverse among players who had a high batting average prior to a trade. Also, my placebo test confirms that the same patterns do not emerge for trades that happened between two regular seasons, providing further evidence that performance resets—rather than inherent differences between inter-league and within-league trades—help to explain why the effects of inter-league trades on post-trade performance depend on pre-trade performance. However, whether a player is traded across leagues or within a league is not exogenous. In subsequent studies, I provide a *causal* test of the effects of performance resets by experimentally manipulating the presence of a performance reset.

<p style="text-align:center">**Study 2**</p>

Study 2 is a laboratory experiment that assesses how an individual's past performance alters the effects of a performance reset on his or her future performance.

Method

Participants. I recruited 202 participants[34] (41.71% female, three unspecified; $M_{age} = 34.49$, three unspecified) from Amazon's Mechanical Turk to take part in a 15-

---

[34] Among the 212 participants who passed my comprehension check questions, I excluded participants who quit the survey *before* being exposed to my experimental manipulations ($n = 9$) or who did not play any games due to technical difficulties ($n = 1$). To conduct an intention-to-treat analysis, I included one participant who quit the game after being exposed to my experimental manipulations and assigned a score

minute study. They were told that in addition to a base pay of $1, they would receive a bonus based on their performance on a word task.

Procedure. Participants were informed that the study involved 10 one-minute Boggle games and that they would receive feedback on their performance after every game. After reading the game rules, participants answered three comprehension-check questions that assessed their understanding of the rules. Participants who failed the comprehension check exited the study and were paid at the base rate, whereas participants who successfully passed the comprehension check proceeded to a one-minute practice game and became the study sample.

For each of the 10 games (plus the practice game), participants were presented with a 3x3 grid of nine letters and were asked to search for as many words as possible for one minute. They were instructed to search for and enter words formed from letters that adjoined horizontally, vertically, or diagonally to the left and right or up and down. According to the game rules, no letter could be used more than once within a single word, and words could not be shorter than three letters.

After each of the first five games, all participants received performance feedback in the same manner. Specifically, a bar graph that marked Games 1-10 on the x-axis depicted the number of correct words participants generated each game. Games that were yet to come had a score of zero. A horizontal line across the bar graph indicated the average number of correct words per game achieved by pretest participants (n = 21).

---

of zero to his unfinished games. However, my results were robust to excluding this participant. This participant did not answer my manipulation check question at the end of the survey. Thus, my analysis of the manipulation check question was conducted based on responses of 201 participants.

Participants were also informed of their average score (i.e., the average number of correct words generated per game) across all of the games they already completed. See Supplemental Materials for complete study materials.

After viewing their performance following Game 5, participants were randomly assigned to either the *performance reset condition* or the *control condition*. In the control condition, participants read, "You have finished five games. Game 6 will begin in one minute. We will continue tracking your average score when the games resume." After a one-minute break, the games resumed, and participants received feedback in the same manner they had in the first five games. Specifically, the bar graph presented after Games 6-10 continued displaying participants' scores for each game, and participants were informed of their running average score, tracked from Game 1.

In the performance reset condition, 10 games were divided into two rounds of five games, and scores achieved in the first round of five games were not included in the tabulation of performance in the second round of five games. Specifically, participants read, "You have finished the first round of five games. Your average score on Round 1 has been saved. A new round will begin in one minute. We will track your average score on Round 2 from zero, a new starting point." Participants were then presented with a new performance interface designed to only track scores for the next five games. After a one-minute break, the games resumed; in other words, Round 2 began. After each game, the bar graph updated participants' score, and participants were informed of their Round 2 running average score, which was tracked from the first game of Round 2. In addition,

participants' average score across the first five games was a constant number and was always presented as part of the feedback interface.

After participants finished with 10 games, they rated the extent to which the five games before the break felt separated from the five games after the break on a seven-point Likert scale (1 = not at all; 7 = very much) and reported their gender and age.

Results

Manipulation check. The effectiveness of the performance reset manipulation was confirmed: the five games before the break felt more separated from the five games after the break for participants in the performance reset condition ($M = 3.53$, $SD = 1.46$) than for those in the control condition ($M = 3.11$, $SD = 1.50$), $t(199) = 2.00$, $p = .046$.

Performance resets and past performance. I measured each participant's *pre-break performance* and *post-break performance* as the number of correct Boggle words generated in the first five games ($M = 35.09$, $SD = 15.45$) and last five games ($M = 37.82$, $SD = 15.00$), respectively. Figure 2 depicts the average difference between pre-break performance and post-break performance as a function of (a) whether a player was in the performance reset or the control condition and (b) whether the player's pre-break performance was below or above the mean. For players whose pre-break performance was below the mean, the improvement between pre-break performance and post-break performance was directionally larger in the performance reset condition ($M_{\text{difference in performance}} = 5.29$, $SD = 7.14$) than in the control condition ($M_{\text{difference in performance}} = 3.23$, $SD = 7.89$), $t(114)$, $p = .14$. However, for players whose pre-break performance was above the mean, the difference between pre-break and post-break performance was significantly

smaller in the performance reset condition ($M_{\text{difference in performance}} = -1.60$, $SD = 10.44$) than in the control condition ($M_{\text{difference in performance}} = 2.59$, $SD = 7.91$), $t(84)$, $p = .04$.

In an ordinary least squares (OLS) regression, I predicted post-break performance using an indicator for the performance reset condition (versus the control conditions), a continuous measure of pre-break performance, and their interaction (performance reset indicator x pre-break performance). The negative and significant interaction between the performance reset indicator and pre-break performance (Panel A of Table 3) suggests that the performance reset treatment influenced post-break performance differently depending on pre-break performance ($p = 0.01$). I conducted a simple slope analysis following the steps outlined by Aiken and West (1991). Compared with the control, a performance reset marginally significantly improved participants' post-break performance for participants whose pre-break performance was one standard deviation below the mean (a gain of 2.65 words; $p = 0.10$); however, a performance reset led to lower post-break performance for participants whose pre-break performance was one standard deviation above the mean (a loss of 3.57 words; $p = 0.03$).

Discussion

In support of Hypothesis 1, Study 2 shows that as compared with the control condition, a performance reset improves post-break performance for participants with poor pre-break performance but harms post-break performance for participants with strong pre-break performance.

Individuals may assess their past performance by comparing their performance with an objective benchmark (if there is any) or with their peers' performance. In the

context of professional baseball players, well-known benchmarks are used to categorize good versus bad performance. In Study 2, participants were provided with their peers' average performance as a benchmark to assess their own performance. Thus, in Studies 1 and 2, I used players' actual past performance to capture whether they viewed their past performance as poor versus strong. In the next two studies, I test Hypothesis 1 differently by experimentally manipulating individuals' perceptions of their past performance. This approach also addresses the concern that differing past performance in Studies 1 and 2 was endogenous.

Study 3

Study 3 extends Study 2 in two ways. First, Study 3 directly manipulated individuals' self-assessment of their past performance instead of measuring past performance. Second, instead of using Boggle games that primarily require cognitive skills, I employed a different task that requires motor skills and responsiveness to changes in stimuli.

Method

Participants. I recruited 244 participants[35] (47.28% female, two unspecified; M_{age} = 32.54, two unspecified) from Amazon's Mechanical Turk to take part in a 15-minute study. They were told that in addition to their base pay of $1, they would receive a bonus based on their performance on a series of games.

---

[35] Among the 260 participants who responded to the survey, I excluded participants who quit the survey before being exposed to my experimental manipulations (n = 14) or whose scores were not successfully registered due to technical problems (n = 2).

Procedure. Participants were informed that the study involved 10 one-minute Snake games (described below) and that they would receive feedback on their performance after every game. They were also told that they would be paid $0.01 for every point that they earned across the 10 games.

To familiarize themselves with the rules of Snake games, participants first played two one-minute practice games. In each game, a white square resembling a snake first appeared at the center of the screen, and a red dot resembling a berry appeared in a random place on the screen. Participants were required to use the arrow keys to move the white square across the screen to touch the red dot. When the white square touched the red dot, the white square increased in length (simulating a snake growing longer after eating a berry), and participants earned one point. Then a new red dot appeared in another random place, waiting for the "snake" to move there. The game would end if (a) the "snake" hit itself while moving around, (b) the "snake" hit the edges of the screen but did not quickly turn around, or (c) the time ran out. This game was designed based on the popular arcade game *Snake*, which was familiar to most of the participants and had a straightforward scoring system.

Following two practice games, participants played 10 real games of Snake and received performance feedback after each game. Similar to Study 2, Study 3 presented participants with their game-by-game scores in a bar graph together with their running average scores. However, unlike Study 2, Study 3 did not include the average score of previous participants as part of the feedback interface because information regarding previous participants' performance was experimentally manipulated (explained below).

See Supplemental Materials for complete study materials. All of the participants played the first five games under the same circumstances. It was only after Game 5 that participants were randomly assigned to one condition of a 2 (*high self-assessment* vs. *low self-assessment*) x 2 (*performance reset* vs. *control*) between-subjects design.

High/low self-assessment manipulation. Prior to Study 3, I recruited 80 individuals from Amazon's Mechanical Turk to complete the same Snake games. I used their performance to set high and low performance standards for the self-assessment experimental manipulation in Study 3. Specifically, participants in Study 3 were told that the program had selected 20 participants who previously worked on this project to help them assess their performance. They were then presented with the average score each of 20 selected pretest participants earned during the first five games. The average scores earned by the 20 pretest participants used in the high self-assessment condition were ranked between the 5th and 25th percentiles of the performance distribution (i.e., from 0.2 to 4.4 points per game). The average scores earned by the 20 pretest participants used in the low self-assessment condition were ranked between the 75th and 95th percentiles of the performance distribution (i.e., from 13.8 to 18 points per game). Thus, the median performance of the 20 pretest participants in the high and low self-assessment conditions was in the 15th and 85th percentiles, respectively. The choice to use pretest scores around the 15th and 85th percentiles of the distribution to manipulate performance expectations is consistent with past research (e.g., Flynn & Amanatullah, 2012).

Performance reset manipulation. The performance reset manipulation in Study 3 was identical to that employed in Study 2. Specifically, in the control condition, the same

bar graph used to record scores for the first five games continued displaying scores for the last five games, and participants' running average score was tracked from Game 1. However, in the performance reset condition, ten games were divided into two rounds of five games. In place of the bar graph used to track scores for the first five games, a new bar graph appeared and only tracked participants' game-by-game scores in the last five games (or Round 2). Participants' Round 2 running average score was tracked from the first game of Round 2, and their Round 1 average score was also included as part of the feedback interface.

After a 30-second break, participants continued to play the last five games. At the end, participants rated the extent to which (a) the five games before the break felt distinct from the five games after the break and (b) the first game after the break felt like a new start (1 = not at all; 7 = very much). They also rated their performance during the first five games (1 = very bad; 7 = very good) and reported their gender and age.

Results

      Manipulation check. The effectiveness of the performance reset manipulation was confirmed: Participants in the performance reset condition reported that (a) the five games before the break felt more distinct from the five games after the break ($M_{distinct}$ = 2.90, SD = 2.04) and (b) the first game after the break felt more like a new start (M_{new_start} = 3.78, SD = 2.13), as compared with participants in the control condition ($M_{distinct}$ = 2.27, SD = 1.81, $t(242)$ = 2.55, p = .01; M_{new_start} = 3.19, SD = 2.09, $t(242)$ = 2.18, p = .03). Also, participants in the high self-assessment condition rated their pre-break performance more positively (M = 4.75, SD = 1.82) than did participants in the low self-assessment

condition ($M = 3.60$, $SD = 1.87$), $t(242) = 4.90$, $p < .0001$, which confirmed the

effectiveness of my self-assessment manipulations.

Performance resets and self-assessments of past performance. I measured

each participant's *pre-break performance* and *post-break performance* as the number of

points earned in the first five games ($M = 45.73$, $SD = 29.65$) and last five games ($M =

52.03$, $SD = 30.31$), respectively. In an OLS regression, I predicted post-break

performance using an indicator for the performance reset condition (versus the control

condition), an indicator for the high (versus low) self-assessment condition, and their

interaction (performance reset indicator x high self-assessment indicator). Following

Flynn and Amanatullah (2012), I included pre-break performance as a control for

individuals' baseline performance. The marginally significant and negative interaction

between the performance reset indicator and the high self-assessment indicator ($p = 0.05$;

Panel B of Table 3) suggests that the effects of a performance reset on post-break

performance depended on individuals' evaluations of their pre-break performance.

Specifically, the planned contrast analysis showed that as compared with the control

condition, a performance reset directionally increased post-break performance (a gain of

0.84 points; $p = 0.78$) for participants in the low self-assessment condition. However, for

participants in the high self-assessment condition, a performance reset significantly

reduced post-break performance (a loss of 7.68 points; $p = 0.01$).

Discussion

Study 3 confirms Hypothesis 1, showing that the effects of a performance reset on

subsequent performance are contingent on individuals' perceptions of their past

performance. Specifically, a performance reset decreases performance when people are led to think positively of their past performance, but (directionally) improves performance when participants are led to view their past performance as poor.

<p style="text-align:center">Study 4</p>

Building on and extending Study 3, Study 4 examines the psychological processes underlying the different effects of performance resets on those who believe they have performed well in the past versus poorly.

Method

Participants. I recruited 81 participants[36] (46.91% female; $M_{age} = 31.28$) from Amazon's Mechanical Turk to take part in a 15-minute study. They were told that in addition to receiving base pay of $1.50, they would receive a bonus based on their performance on a series of games.

Procedure. Participants played 10 one-minute Snake games as described in Study 3 and received feedback on their performance after every game. After Game 5, they were randomly assigned to one condition of a 2 (*high self-assessment* vs. *low self-assessment*) x 2 (*performance reset* vs. *control*) between-subjects experiment. All materials and measures in Study 4 were identical to those used in Study 3, with two exceptions. First, I collected measures of potential mechanisms after 10 games were completed (described below). Second, I adapted the benchmarks used to manipulate perceptions of past performance. As noted above, participants in the low self-assessment

---

[36] Among the 106 participants who responded to my survey, I excluded participants who quit the survey before being exposed to my experimental manipulations ($n = 24$) or who were unable to play the games due to technical difficulties ($n = 1$).

condition of Study 3 only directionally benefited from a performance reset (compared to the control). One possibility is that the high standards set by pretest players (i.e., from the 75th to the 95th percentile of pretest performance) were so daunting that a performance reset was unable to boost the morale of participants in the low self-assessment condition. Thus, in Study 4, I selected previous players whose performance would set a more reasonable benchmark. Specifically, the low self-assessment condition relied on 20 players whose average scores on the first five games were between the 65th and 95th percentiles in Study 3 (i.e., from 12.6 to 17 points per game). The high self-assessment condition relied on 20 players whose average scores on the first five games were between the 5th and 15th percentiles in Study 3 (i.e., from 0.4 to 1.2 points per game).

Mechanism measures

Self-efficacy. I assessed self-efficacy with the confidence scores developed by Bandura (2006). Specifically, I asked participants to recall how certain they were right before starting the last five games that they could (a) perform better in the last five games than in the first five games and (b) improve their performance rank relative to other participants in the last five games. Participants rated their degree of confidence by selecting a number from 0 to 100 using a scale adapted from Bandura (2006): 0 = Could not do at all; 50 = Moderately certain could to; 100 = Highly certain could do. The variable, *self-efficacy*, equaled the summation of participants' confidence ratings for the two items (correlation coefficient = 0.79) following Brown et al. (2005).

Commitment. I used two measures to capture individuals' commitment to performing well during the last five games. First, participants indicated how committed

they were to performing well in the last five games, relative to the first five games (-3 = Much less committed in Games 6-10[/Round 2]; 0 = About the same in Games 6-10[/Round 2]; 3 = Much more committed in Games 6-10[/Round2]).[37] This measure was adapted from the target-free assessment of commitment developed by Klein, Cooper, Molloy, and Swanson (2014). In addition, participants rated how important it was for them to perform well in the last five games (relative to the first five games) in order to show that they were good at playing the Snake games (-3 = Much less important in Games 6-10[/Round 2]; 0 = About the same in Games 6-10[/Round 2]; 3 = Much more important in Games 6-10[/Round2]). This measure captured the perceived value of achieving strong performance, a determinant of commitment. The variable, *commitment*, equaled the average of a participant's responses to these two items (correlation coefficient = 0.76).

Additional measures (positive and negative affect) were collected at the end to address the possibility that affect explained the effects of performance resets and past performance on future performance. Neither positive affect nor negative affect varied by condition. Details about the measures and results are reported in Appendix F.

Results

Manipulation check. The effectiveness of the performance reset manipulation was confirmed: Participants in the performance reset condition reported that (a) the five games before the break felt more distinct from the five games after the break ($M_{distinct}$ = 2.53, SD = 1.58) and (b) the first game after the break felt more like a new start (M_{new_start}

---

[37] The wording in the brackets was used in the performance reset condition.

= 3.30, *SD* = 1.94) than did participants in the control condition ($M_{distinct}$ = 3.44, *SD* =

1.86, *t*(79) = 2.38, *p* = .02; M_{new_start} = 4.83, *SD* = 1.96, *t*(79) = 3.53, *p* < .001). The self-

assessment manipulation also successfully affected participations' evaluations of their

pre-break performance: participants in the high self-assessment condition rated their pre-

break performance more positively (*M* = 5.05, *SD* = 1.53) than did participants in the low

self-assessment condition (*M* = 2.95, *SD* = 1.89), *t*(79) = 5.51, *p* < .0001.

        Performance resets and self-assessments of past performance. As in Study 3, I

measured each participant's *pre-break performance* and *post-break performance* as the

number of points earned in the first five games (*M* = 46.84, *SD* = 27.82) and last five

games (*M* = 52.36, *SD* = 29.25), respectively. In an OLS regression, I predicted post-

break performance using an indicator for the performance reset condition (versus the

control condition), an indicator for the high (versus low) self-assessment condition, their

interaction (performance reset indicator x high self-assessment indicator), and a control

for pre-break performance. A significant and negative interaction between the

performance reset indicator and the high self-assessment indicator (*p* = 0.03; Model 1 of

Table 4) suggests that the effects of a performance reset on post-break performance

depend on individuals' perceptions of their pre-break performance. The planned contrast

analysis shows that a performance reset directionally increased post-break performance (a

gain of 1.69 points; *p* = 0.72) for participants in the low self-assessment condition but

significantly reduced post-break performance for participants in the high self-assessment

condition (a loss of 12.52 points; *p* = 0.006).

146

Mechanism analysis. Hypothesis 2 predicts that performance resets are more likely to boost self-efficacy for individuals who view their past performance as weak than for individuals who think positively of their past performance. To test Hypothesis 2, I predicted self-efficacy using the performance reset indicator, the high self-assessment indicator, and their interaction. The significant and negative interaction between the reset indicator and the high self-assessment indicator ($p = 0.049$; Model 2 in Table 4) suggests that the effects of a performance reset on self-efficacy hinged on individuals' assessments of their past performance. The planned contrast analysis shows that for participants in the low self-assessment condition, a performance reset increased self-efficacy as compared to the control condition ($p = .001$); however, for participants in the high self-assessment condition, a performance reset did not significantly change self-efficacy ($p = .57$).

Hypothesis 3 predicts that past performance would affect the effects of performance resets on individuals' commitment to performing well. To test Hypothesis 3, I predicted commitment using the reset indicator, the high self-assessment indicator, and their interaction. The interaction between the reset indicator and the high self-assessment indicator is not statistically significant ($p = 0.94$; Model 3 in Table 4), suggesting that the effects of a performance reset on commitment did not depend on individuals' assessments of their pre-break performance. In fact, the performance reset treatment did not significantly change commitment in either the high self-assessment condition ($p = 0.75$) or the low self-assessment condition ($p = .82$). Thus, Hypothesis 3 was not supported.

I next tested the mediated-moderation model in which self-efficacy and commitment were expected to mediate the relationship between a performance reset,

perceptions of pre-break performance, and post-break performance (Hypothesis 4).

Following Preacher, Rucker, and Hayes (2007) and Preacher and Hayes (2008), I

simultaneously tested whether self-efficacy and commitment were significant mediators.

First, as described above, I have established that self-assessments of past performance

moderate the relationship between a reset and self-efficacy (p = .049; Model 2 in Table

4), but do not moderate the relationship between a reset and commitment (p = .82; Model

3 in Table 4). Next, I conducted a 5,000-sample bootstrap analysis to estimate the indirect

effects of a performance reset via self-efficacy and commitment when self-assessments of

past performance were manipulated to be high versus low. [38] As reported in the bottom

half of Table 4, the indirect effect of a performance reset on post-break performance via

self-efficacy is statistically significant in the low self-assessment condition (B = 4.05, SE

= 2.43, 95% bias-corrected confidence interval excluding zero [0.50, 10.46]); however,

the indirect effect is not significant in the high self-assessment condition (B = 0.67, SE =

1.27, 95% bias-corrected confidence interval [-1.36, 4.18]). This suggests that self-

efficacy is a significant mediator of the process through which self-assessments of past

performance influence the effects of a performance reset on future performance.

However, I did not find evidence that commitment was a significant mediator of the

relationship between perceptions of past performance, a reset, and future performance.

Specifically the indirect effects of a reset on post-break performance via commitment

were not statistically significant in the low self-assessment condition (B = -0.03, SE =

---

[38] Similar to other methodologists (e.g., Edwards & Lambert, 2007; Muller, Judd, & Yzerbyt, 2005),
Preacher et al., (2007) and Preacher and Hayes (2008) recommended directly examining the indirect effects
using a product of coefficients approach together with a bootstrap analysis.

0.73, 95% CI [-1.80, 1.29]) or the high self-assessment condition (B = -0.02, SE = 0.83, 95% CI [-2.24, 1.31]). Altogether, Hypothesis 4 is only partly supported.

Discussion

Study 4 replicates the results in Study 3, showing that a performance reset (directionally) improves performance when individuals are led to perceive their past performance as poor but significantly harms performance when individuals are led to perceive their past performance as strong. Study 4 also provides preliminary evidence that perceptions of past performance moderate the effects of a performance reset on self-efficacy, which further influences performance. Further, Study 4 fails to support that commitment to performing well is one reason why self-assessments of past performance alter the effects of a performance reset on future performance.

Study 5

Thus far, I have shown that whether a performance reset improves or dampens future performance hinges on how individuals view their past performance. Study 5 offered individuals an opportunity to reset their performance statistics. Since participants were incentivized to perform well, Study 5 was designed to test my theory in a new way by examining whether people would rationally differentiate the value of performance resets as a function of their perceptions of past performance.

Method

Participants. I recruited 183 participants[39] (63.39% female; $M_{age} = 20.71$) from a northeastern university in the United States. This experiment was part of a series of studies participants completed during a one-hour laboratory session. Participants were told that in addition to receiving a flat participation fee, they would receive a bonus based on their performance on a word task.

Procedure. Study 5 used the same Boggle games employed in Study 2. Participants were first informed that the study involved 10 one-minute Boggle games and that they would receive performance feedback after every game. After passing comprehension check questions and playing a practice game, participants proceeded to the games that counted. The remaining procedure in Study 5 was the same as the procedure in Study 3 with one key difference: Participants in Study 5 were asked to choose whether or not to reset their performance after Game 5 rather than being randomly assigned to either a performance reset or a control condition.

As in Study 3, participants in Study 5 first played five games without receiving information about their peers' performance and were randomly assigned to one of two experiment conditions after Game 5: *high self-assessment condition* or *low self-assessment condition*.

High/low self-assessment manipulation. Prior to conducting Study 5, I recruited 178 participants from the same sample used in Study 5 to complete five, one-minute Boggle games in addition to a practice game. Participants in the high self-assessment

---

[39] Among the 196 participants who took the survey, I excluded participants who could not play the games due to technical difficulties ($n = 1$) or who took the survey twice ($n = 12$). My results were robust if I included every participant who responded to my dependent measure (described later) for the first time.

condition of Study 5 were presented with the average scores of 20 pretest participants whose scores ranked between the 5th and 15th percentiles of the performance distribution (i.e., from 2.4 to 3.6 correct words per game); participants in the low self-assessment condition were presented with the average scores of 20 pretest participants whose scores ranked between the 85th and 95th percentiles of the performance distribution (i.e., from 12.8 to 18.2 correct words per game). Participants in Study 5 were asked to write down how many of these 20 previous participants had an average score lower than their own average score (see Supplemental Materials for complete study materials).

Participants subsequently were told that they had two different options for the next five games to be played. One option was that they could play the next five games as Games 1-5 of a new round, in which case their average score on the first round of five games would be saved and their average score on the second round would be tracked from zero. The other option was that they could continue to play the next five games as Games 6-10 without starting a new round. For each option, I showed participants a graph depicting what their performance interface would look like if they ended up choosing that option. I counterbalanced the order in which the two options were described (which did not affect participants' decision). After participants indicated their choice, they played five more games and received feedback in accordance with their choice.

After completing 10 games, participants rated their performance during the first five games (1 = very bad; 7 = very good) and were asked to speculate about the purpose of this study. The results reported below are robust if I exclude the 21 participants who

correctly guessed the purpose of the study. Last, participants reported their gender and age and were paid according to the total number of correct words they generated.

Results

  Manipulation check. I first confirmed that participants in the high self-assessment condition rated their performance on Games 1-5 more positively ($M = 5.47$, $SD = 1.14$) than did participants in the low self-assessment condition ($M = 4.14$, $SD = 1.33$), $t(181) = 7.23$, $p < .0001$.

  The decision to start a new round. The primary measure of interest was whether or not individuals chose to start a new round. Participants in the low self-assessment condition chose to start a new round and reset their average scores at a higher rate than participants in the high self-assessment condition (52.75% vs. 34.78%), $\chi^2(1) = 6.00$, $p = .01$. Ancillary analyses show that compared with those in the high self-assessment condition, participants in the low self-assessment condition were more likely to expect resetting performance statistics to improve their subsequent performance and thus elect to reset their statistics at a higher rate (see Appendix G for details).

Discussion

  Participants in Study 5 were paid based on their performance and thus had the incentive to choose an option that they believed would help them earn more points. I find that relative to those who think more positively of their past performance, people with a less positive view are more likely to choose to wipe their slate clean. Combining this result with earlier findings that resets improve performance for those with a negative view of their past performance but harm performance for those with a positive view of

their past performance, Study 5 suggests that people rationally differentiate the value of resetting their statistics based on how they perceive their past performance.

General Discussion

Across five field and lab studies, I demonstrated that the effects of performance resets on future performance depend on past performance. Specifically, performance resets improve future performance for individuals who view their recent performance as weak but harm future performance for individuals who view their recent performance as strong. These patterns held when I used objective performance to capture how individuals would assess their past performance (Studies 1 & 2) and when I experimentally manipulated individuals' perceptions of their past performance (Studies 3 & 4). Study 4 provided initial support for a mediated-moderation model in which a performance reset boosted self-efficacy to a greater extent among those with undesirable recent performance (compared with those who had performed well), and self-efficacy further positively affected future performance. However, Study 4 failed to support the hypothesis that commitment mediated the relationship between past performance, a reset, and future performance. Further, Study 5 showed that perceptions of past performance shift preferences for performance resets, whereby individuals chose to reset their performance statistics more frequently when led to believe their past performance was poor than when led to believe it was strong.

Theoretical Implications

Management research that examines the relationship between past performance and future performance has largely assumed that past performance is related to who we

are now and thus alters our self-efficacy and commitment (e.g., Bandura, 1991, 1997; Fishbach & Dhar, 2005). However, there has been little research on how an individual's psychological separation from her past successes and failures may affect her future motivation and effort allocation. Extending recent research showing that landmark events on calendars and in personal histories create fresh starts and increase goal pursuit, this paper provides a theoretical framework that predicts when fresh starts on performance records induced by performance resets can be inspiring versus demotivating. In addition, this paper examines changes in actual performance rather than the initiation of goal-pursuit progress (Dai et al., 2014, 2015) or optimism associated with feeling separated from past failures (Libby & Eibach, 2002).

Several studies in this paper manipulated individuals' self-assessments of their past performance by providing social-comparison information. Information about peers with strong versus poor performance can be viewed as negative versus positive feedback. Thus, the current research also contributes to the literature on performance feedback by identifying a situational factor that influences whether positive or negative feedback leads to higher subsequent performance. Specifically, the findings in this paper suggest that negative feedback may be more motivating than positive feedback when resetting your performance record is possible, whereas positive feedback may be more motivating than negative feedback when your past performance will continue to be part of performance appraisals. This pattern may arise because the combination of a performance reset and negative feedback facilitates people's tendency to disparage their imperfect past selves (Ross & Wilson, 2002), leading them to think more positively of their current selves and

gain higher confidence. Future research that directly tests whether it is easier to internalize and leverage negative feedback after a fresh start would be valuable.

Practical Implications

The results documented in this paper have a number of practical implications for managers. First, when deciding whether to reset employees' performance metrics to zero, they should recognize that performance resets may affect employees differently depending on their past performance (which may be measured on an objective scale or relative to their coworkers). In particular, managers need to be aware of the negative impact of performance resets on high performers when they plan to implement an intervention that will disregard employees' past records. To reduce the possibility that new programs will backfire among top performers, managers may want to highlight the relevance of past performance only for top performers.

Second, this paper, and Study 5 in particular, suggest that employees would appreciate the opportunity to put their past performance records behind them after a streak of weak performance. Managers may be able to help employees cope with negative feedback by offering them the option to reset their performance statistics (e.g., letting employees with low daily sales the previous week choose whether they want to continue tracking their daily sales over the entire month or only over the course of the current week). Further, employees endowed with such an option may feel control over how they can motivate themselves and thus experience increased job satisfaction and engagement.

Limitations and Future Directions

An important question related to the practical implications of performance resets is how long the patterns observed in this paper persist. One limitation of my laboratory experiments is that I was only able to track performance over a short period following a performance reset. However, Study 1 showed that the benefits of performance resets in the MLB lasted for up to two months. Given the robust evidence across Studies 2-4 that performance resets dampen high performers' subsequent performance, my findings have meaningful implications even if performance resets only have short-term effects. However, it would valuable for future research to examine the long-term effects of performance resets and identify the right time for performance resets to recur in order to maximize the motivating impact of resets for low performers.

Studies 2-5 examine the effects of performance resets that individuals do not anticipate.[40] Not informing participants of the existence of a performance reset allowed me to keep information and expectations the same across conditions before my performance reset treatment. This type of unexpected performance reset has parallels in organizations. For example, a manager may choose not to announce a new "employee of the month" program in advance, leaving employees unaware that their past records will not be counted toward the new program until the beginning of a new month. Future studies could explore how employees anticipating a performance reset adjust their efforts *before* the reset occurs, a question that will help managers decide whether and when to reveal interventions that may (inadvertently) reset employees' performance records.

---

[40] In Study 1, whether an MLB player knows in advance that he will be traded depends on many factors, such as whether he is to be traded after the MLB's July 31 trade deadline and how long he has been on an active major league roster.

In addition, my studies focused on performance resets that tracked performance from zero but still counted individuals' past performance toward their final rewards (e.g., professional players' career statistics, the bonus participants received at the end of an experiment). A stronger form of performance reset may completely disregard past records. For example, when an employee switches to a different branch office, her promotion prospects may be determined solely by her performance in the new position. Future research exploring and comparing a broad set of operationalizations of performance resets would be valuable.

Conclusion

Given the prevalence of work contexts where managers intentionally or inadvertently disregard employees' past records when tabulating and rewarding their future performance, the need to understand the relationship between performance resets and individual performance remains critical. Specifically, does wiping one's performance record clean enhance or impair individual performance? Five studies show that the introduction of a performance reset can produce performance benefits for individuals who did not perform well lately but cause performance decrements for those who have been effective in the past. I proposed that self-efficacy and commitment—two important determinants of individual performance—may drive the psychological processes underlying these results. Specifically, performance resets may improve performance by enhancing self-efficacy and commitment when past performance is viewed as weak but harm performance by decreasing commitment without improving self-efficacy when past performance is viewed as strong. By recognizing that not all employees are affected in

the same manner, organizations and their members may be able to better harness the

benefits of and avoid the harms of performance resets.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287. doi: 10.1016/0749-5978(91)90022-L

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Charlotte, NC: Information Age.

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87-99. doi: 10.1037/0021-9010.88.1.87

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586-598. doi: 10.1037/0022-3514.41.3.586

Bartels, D. M., & Urminsky, O. (2011). On intertemporal selfishness: How the perceived instability of identity underlies impatient consumption. *Journal of Consumer Research, 38*(1), 182-198. doi: 10.1086/658339

Baseball Almanac. (n.d.). Offensive stats 101. Retrieved March 17, 2015, from http://www.baseball-almanac.com/stats.shtml

Brown, S. P., Jones, E., & Leigh, T. W. (2005). The attenuating effect of role overload on relationships linking self-efficacy and goal level to work performance. *Journal of Applied Psychology, 90*(5), 972-979. doi: 10.1037/0021-9010.90.5.972

Cochran, W., & Tesser, A. (1996). The "what the hell" effect: Some effects of goal proximity and goal framing on performance. In L. L. Martin & A. Tesser (Eds.), *Striving and feeling: Interactions among goals, affect, and self- regulation*. Hillsdale, NJ: Psychology Press

Dai, H., Milkman, K. L., & Riis, J. (2014). The fresh start effect: Temporal landmarks motivate aspirational behavior. *Management Science, 60*(10), 2563-2582. doi: 10.1287/mnsc.2014.1901

Dai, H., Milkman, K. L., & Riis, J. (2015). *Put your imperfections behind you: Why and how meaningful temporal landmarks motivate aspirational behavior*. Working Paper. Retrieved from http://ssrn.com/abstract=2420476

Donovan, J. J., & Williams, K. J. (2003). Missing the mark: Effects of time and causal attributions on goal revision in response to goal-performance discrepancies. *Journal of Applied Psychology, 88*(3), 379-390. doi: 10.1037/0021-9010.88.3.379

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods, 12*(1), 1-22. doi: 10.1037/1082-989x.12.1.1

Fishbach, A., & Dhar, R. (2005). Goals as excuses or guides: The liberating effect of perceived goal progress on choice. *Journal of Consumer Research, 32*(3), 370-377. doi: 10.1086/497548

Flynn, F. J., & Amanatullah, E. T. (2012). Psyched up or psyched out? The influence of coactor status on individual performance. *Organization Science, 23*(2), 402-415. doi: 10.1287/orsc.1100.0552

Gubler, T., Larkin, I., & Pierce, L. (2014). *Motivational spillovers from awards: Crowding out in a multitasking environment*. Working Paper. Washington University in St. Louis. St Louis, MO.

Heggestad, E. D., & Kanfer, R. (2005). The predictive validity of self-efficacy in training performance: Little more than past performance. *Journal of Experimental Psychology-Applied, 11*(2), 84-97. doi: 10.1037/1076-898x.11.2.84

Imas, A. (2014). *The realization effect: Risk-taking after realized versus paper losses*. Working Paper. Carnegie Mellon University. Pittsburgh, PA.

Klein, H. J., Cooper, J. T., Molloy, J. C., & Swanson, J. A. (2014). The assessment of commitment: Advantages of a unidimensional, target-free approach. *Journal of Applied Psychology, 99*(2), 222-238. doi: 10.1037/A0034751

Klein, H. J., & Kim, J. S. (1998). A field study of the influence of situational constraints, leader-member exchange, and goal commitment on performance. *Academy of Management Journal, 41*(1), 88-95. doi: 10.2307/256900

Libby, L. K., & Eibach, R. P. (2002). Looking back in time: Self-concept change affects visual perspective in autobiographical memory. *Journal of Personality and Social Psychology, 82*(2), 167-179. doi: 10.1037//0022-3514.82.2.167

Locke, E. A., Frederick, E., Lee, C., & Bobko, P. (1984). Effect of self-efficacy, goals, and task strategies on task-performance. *Journal of Applied Psychology, 69*(2), 241-251. doi: 10.1037/0021-9010.69.2.241

Locke, E. A., Latham, G. P., & Erez, M. (1988). The determinants of goal commitment. *Academy of Management Review, 13*(1), 23-39. doi: 10.2307/258352

Mandelbaum, M. (2005). *The meaning of sports*. New York: PublicAffairs.

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*(6), 852-863. doi: 10.1037/0022-3514.89.6.852

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes - a meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30-38. doi: 10.1037/0022-0167.38.1.30

Peetz, J., & Wilson, A. E. (2013). The post-birthday world: Consequences of temporal landmarks for temporal self-appraisal and motivation. *Journal of Personality and Social Psychology, 104*(2), 249-266. doi: 10.1037/A0030477

Polivy, J., & Herman, C. P. (2002). If at first you don't succeed: False hopes of self-change. *American Psychologist, 57*(9), 677-689. doi: 10.1037//0003-066x.57.9.677

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavioral Research Methods, 40*(3), 879-891. doi: 10.3758/BRM.40.3.879

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research, 42*(1), 185-227.

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty, 19*(1-3), 171-197. doi: 10.1023/A:1007879411489

Retrosheet. (n.d.). Play-by-play data files (event files). Retrieved December 20, 2014, from http://www.retrosheet.org/game.htm

Riedel, J. A., Nebeker, D. M., & Cooper, B. L. (1988). The influence of monetary incentives on goal choice, goal commitment, and task-performance. *Organizational Behavior and Human Decision Processes, 42*(2), 155-180. doi: 10.1016/0749-5978(88)90010-6

Ross, M., & Wilson, A. E. (2002). It feels like yesterday: Self-esteem, valence of personal past experiences, and judgments of subjective distance. *Journal of Personality and Social Psychology, 82*(5), 792-803. doi: 10.1037//0022-3514.82.5.792

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. doi: 10.1037//0003-066x.55.1.68

Schmidt, A. M., & DeShon, R. P. (2010). The moderating effects of performance ambiguity on the relationship between self-efficacy and performance. *Journal of Applied Psychology, 95*(3), 572-581. doi: 10.1037/A0018289

Schunk, D. H., & Hanson, A. R. (1985). Peer models - influence on childrens self-efficacy and achievement. *Journal of Educational Psychology, 77*(3), 313-322. doi: 10.1037//0022-0663.77.3.313

Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer-model attributes and childrens achievement behaviors. *Journal of Educational Psychology, 79*(1), 54-61. doi: 10.1037//0022-0663.79.1.54

Schwarz, A. (2007). For pitchers, national league really is an easier out, *The New York Times*. Retrieved from

http://www.nytimes.com/2007/01/14/sports/baseball/14score.html?fta=y&_r=0

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*(2), 240-261. doi: Doi 10.1037/0033-2909.124.2.240

Stefano, G. D., Gino, F., Pisano, G., & Staats, B. (2014). *Learning by thinking: How reflection aids performance*. Harvard Business School Working Paper 14-093.

Sue-Chan, C., & Ong, M. (2002). Goal assignment and performance: Assessing the mediating roles of goal commitment and self-efficacy and the moderating role of power distance. *Organizational Behavior and Human Decision Processes, 89*(2), 1140-1161. doi: 10.1016/S0749-5978(02)00017-1

Tolli, A. P., & Schmidt, A. M. (2008). The role of feedback, causal attributions, and self-efficacy in goal revision. *Journal of Applied Psychology, 93*(3), 692-701. doi: 10.1037/0021-9010.93.3.692

Vancouver, J. B. (1997). The application of HLM to the analysis of the dynamic interaction of environment, person and behavior. *Journal of Management, 23*(6), 795-818.

White, R. W. (1959). Motivation reconsidered - the concept of competence. *Psychological Review, 66*(5), 297-333. doi: 10.1037/H0040934

Wilson, A. E., & Ross, M. (2001). From chump to champ: People's appraisals of their earlier and present selves. *Journal of Personality and Social Psychology, 80*(4), 572-584. doi: 10.1037//0022-3514.80.4.572

Zhang, Y., & Huang, S. C. (2010). How endowed versus earned progress affects consumer goal commitment and motivation. *Journal of Consumer Research, 37*(4), 641-654. doi: 10.1086/655417

Table 1. Summary Statistics and Comparisons Between Two Trade Types (Study 1)

Panel A: Means, standard deviations, and correlations among study variables

Variable	Mean	SD	1	2	3	4
1 Hitting Indicator	0.26	0.44	1.00			
2 Inter-league Indicator	0.42	0.49	0.00	1.00		
3 Pre-trade Batting Average	0.26	0.03	0.05*	-0.01*	1.00	
4 Post-trade Indicator	0.36	0.48	0.01*	-0.02*	-0.03*	1.00

Note. $N = 268{,}491$ at bats

* p-value < .05

Panel B: Inter-league trade and within-league trade comparisons

Variable	Within-league Trades	Across-league Trades	P-value for the t-test
Pre-trade Batting Average	0.256 (0.002)	0.255 (0.002)	0.69
Number of "At Bats" Prior to Trade	240.42 (4.80)	245.09 (5.86)	0.54
Number of "At Bats" Post Trade	157.48 (4.99)	147.79 (5.17)	0.19
Month of Trade (Apr-Sept Indicated by 4-9)	7.12 (0.04)	7.11 (0.05)	0.93
Days between First Pre-Trade Game and Trade	109.1 (1.42)	111.45 (1.52)	0.27
Days between Last Pre-Trade Game and Trade	2.98 (0.32)	3.11 (0.42)	0.79
Days between First Post-Trade Game and Trade	2.25 (0.28)	2.25 (0.35)	0.98
Days between Last Post-Trade Game and Trade	62.24 (1.35)	61.69 (1.50)	0.79

Note. $N = 706$ player-seasons
(within-league trade: 410 player-seasons; across-league trade: 296 player-seasons)

Table 2. Results of Regressions Predicting Probability of Hitting at Bat (Study 1)

Variable	Model 1: Mid-season trades	Model 2: Between-season trades
Post-trade indicator	0.148***	0.165***
	(0.017)	(0.012)
Post-trade indicator X Inter-league indicator	0.064*	0.004
	(0.026)	(0.018)
Post-trade indicator X Pre-trade batting average	-0.557***	-0.626***
	(0.067)	(0.047)
Post-trade indicator X Inter-league indicator X Pre-trade batting average	-0.247*	-0.030
	(0.101)	(0.069)
Observations	268,491	732,415
Number of player-seasons (fixed effects included)	706	1,124
R^2	0.0047	0.0037

Note. Standard errors are clustered at the player-season level.

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Table 3. Results of Regressions Predicting Post-Break Performance (Studies 2 and 3)

Panel A: Study 2

Variable	Model 1
Performance reset indicator	6.61*
	(2.89)
Pre-break performance	0.94***
	(0.06)
Performance reset indicator X Pre-break performance	-0.20**
	(0.08)
Observations	202
R^2	0.7136

^, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

Panel B: Study 3

Variable	Model 1
Performance reset indicator	0.84
	(3.05)
High self-assessment indicator	4.92
	(3.07)
Performance reset indicator X High self-assessment indicator	-8.51*
	(4.33)
Pre-break performance	0.84***
	(0.04)
Observations	244
R^2	0.6951

^, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% levels, respectively.

Table 4. Total Effects of Performance Resets and Self-Assessments of Past Performance on Post-Break Performance and Their Indirect Effects Through Self-Efficacy and Commitment (Study 4)

Variable	Model 1: Post-break performance	Model 2: Self-efficacy	Model 3: Commitment
Performance reset indicator	1.69	57.99***	-0.14
	(4.61)	(17.49)	(0.44)
High self-assessment indicator	3.35	41.29*	-0.38
	(4.57)	(17.34)	(0.44)
Performance reset indicator	-14.21*	-48.38*	0.04
X High self-assessment indicator	(6.37)	(24.22)	(0.62)
Pre-break performance	0.91***		
	(0.06)		
Observations	81	81	81
R^2	0.7791	0.1368	0.0173

Conditional indirect effects	Indirect effect	Bootrapped SE	95% bias corrected confidence interval
Self-efficacy as the mediator: Low self-assessment	4.05	2.43	[0.50, 10.46]
Self-efficacy as the mediator: High self-assessment	0.67	1.27	[-1.36, 4.18]
Commitment as the mediator: Low self-assessment	-0.03	0.73	[-1.80, 1.22]
Commitment as the mediator: High self-assessment	-0.02	0.83	[-2.24, 1.31]

Note. Bootstrap sample size = 5,000

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Figure 1. Major League Baseball Players' Average Probability of Hitting at Bat Before

and After Mid-Season Trades (Study 1)
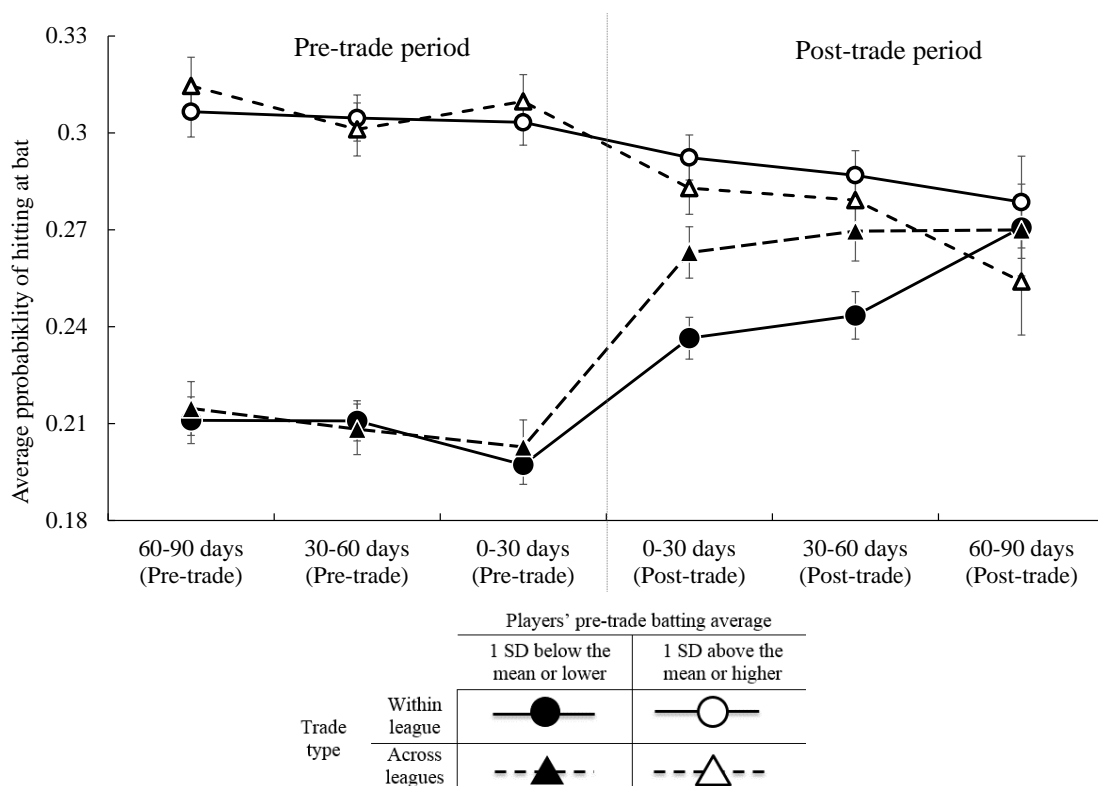
Note. Error bars correspond to standard errors.

Figure 2. Change in Performance Between the First and Last Five Games as a Function of the Presence of a Performance Reset and Pre-Break Performance (Study 2)
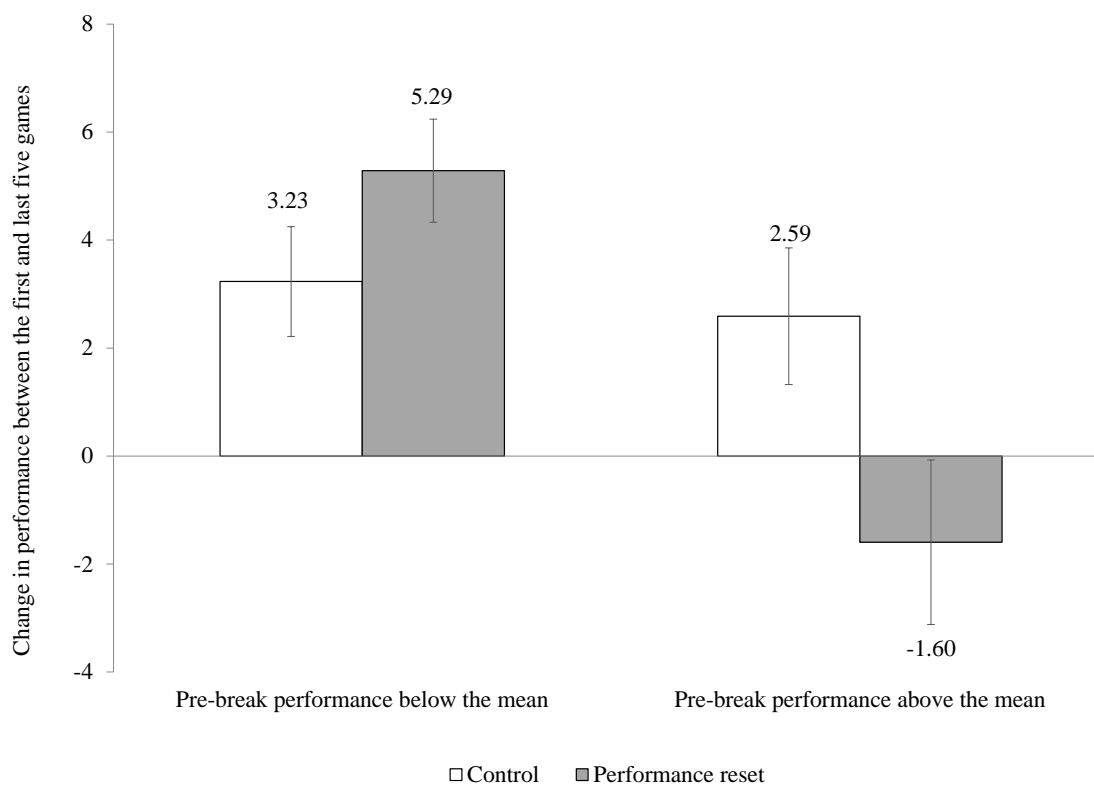
Note. Error bars correspond to standard errors.

Appendix A. Analysis of the Effects of Inter-league Trades on Pitchers (Study 1)

　　　I focused on batters rather than pitchers for the reasons explained in the paper. However, for the sake of being comprehensive, I used the same play-by-play data reported in the paper to analyze how inter-league trades (versus within-league trades) affect pitchers' strikeout rate, one of the most common performance metrics for pitchers. I relied on the same regression specification described in the Analysis Strategy section of the paper, with three changes. First, the dependent variable was a dummy variable indicating whether or not a given pitcher caused a strikeout when he faced the batter in question. Second, I operationalized *past performance* as a given pitcher's strikeout rate (the total number of strikeouts he caused divided by the total number of batters he faced)[41] prior to a trade. Third, I restricted my sample to pitchers who faced at least 100 batters prior to a trade. As shown in the table below, inter-league trades do not significantly improve post-trade performance over and beyond within-league trades (*p* = .81), and pre-trade strikeout rate does not moderate the effects of inter-league trades on post-trade strikeout rate (*p* = .92).

Variable	Model 1
Post-trade indicator	0.049***
	(0.006)
Post-trade indicator X Inter-league indicator	0.003
	(0.010)
Post-trade indicator X Pre-trade strikeout rate	-0.299***
	(0.043)
Post-trade indicator X Inter-league indicator X Pre-trade strikeout rate	0.007
	(0.068)
Observations	369,053
Number of player-seasons (fixed effects included)	776
R^2	0.0174

Note. Standard errors are clustered at the player-season level.

*** denotes significance at the 0.1% level.

[41] FanGraphs. (n.d.). Retrieved March 16, 2015 from http://www.fangraphs.com/library/pitching/rate-stats/

Appendix B. Placebo Tests Using Trades that Happened Between Seasons (Study 1)

Inter-league trades occurring between seasons do not reset performance statistics. If the findings reported in the paper are primarily driven by performance resets that are induced by mid-season inter-league trades, then I should see little to no effect among trades that occurred between seasons. I examined 1,111 cases where a player (a) was traded between two seasons, (b) had at least 100 at bats while he was playing for his previous team in the season prior to the trade, and (c) played at least one game for his new team in the season right after the trade. I used the same regression specification described in the Analysis Strategy section of the paper. As shown in the table below, the interaction between the post-trade indicator and the inter-league indicator is not statistically significant ($p = .82$), suggesting that inter-league trades do not improve post-trade hitting probability over and beyond within-league trades. Also, the interaction between the post-trade indicator, the inter-league indicator, and pre-trade batting average is not statistically significant ($p = .66$), suggesting that pre-trade batting average does not moderate the effects of inter-league trades on post-trade hitting probability.

Variable	Model 1
Post-trade indicator	0.165***
	(0.012)
Post-trade indicator X Inter-league indicator	0.004
	(0.018)
Post-trade indicator X Pre-trade batting average	-0.626***
	(0.047)
Post-trade indicator X Inter-league indicator X Pre-trade batting average	-0.030
	(0.069)
Observations	732,415
Number of player-seasons (fixed effects included)	1,124
R^2	0.0037

Note. Standard errors are clustered at the player-season level.

*** denotes significance at the 0.1% level.

Appendix C. Results of a Regression that Sliced the Post-trade Period into Four Narrower Windows (Study 1)

   In an alternative regression model, I replaced the post-trade indicator with four post-trade indicators, which represented 0-30 days, 30-60 days, 60-90 days, and beyond 90 days after a trade, respectively. Similar to Figure 1 in the paper, this nuanced analysis allows me to capture how players' responses to inter-league trades vary over time. In the table below, *post-trade month 1* was a dummy variable that equaled 1 if a given observation occurred within 30 days after the trade in question and otherwise equaled 0. Similarly, *post-trade month 2*, *post-trade month 3,* and *post-trade month 4* indicated the period of 30-60 days, 60-90 days, and more than 90 days after a trade.

   As shown in the table below, the positive and significant coefficients on the four post-trade indicators (all ps $< .01$) mean that the probability of hitting at bat on average improved after a within-league trade. Further, hitting probability on average improved more substantially after an inter-league trade than after a within-league trade in the first two months following a trade, as indicated by the positive and significant interaction (a) between *post-trade month 1* and the *inter-league indicator* ($p = .028$) and (b) between *post-trade month 2* and the *inter-league indicator* ($p = .042$). Importantly, the three-way interaction between *post-trade month 1,* the *inter-league indicator*, and *pre-trade batting average* ($p = .013$) is negative and significant, as is the interaction between *post-trade month 2,* the *inter-league indicator*, and *pre-trade batting average* ($p = .047$). This suggests that the lower a batter's pre-trade batting average, the more an inter-league trade would improve his post-trade hitting probability above and beyond a within-league trade in the first two months following a trade. After 60 days following a trade, the aforementioned effects were not statistically significant. These findings are consistent with Figure 1 and Hypothesis 1.

Variable	Model 1
Post-trade month 1	0.126***
	(0.023)
Post-trade month 2	0.140***
	(0.025)
Post-trade month 3	0.245***
	(0.041)
Post-trade month 4	0.163**
	(0.060)
Post-trade month 1 X Inter-league indicator	0.073*
	(0.033)
Post-trade month 2 X Inter-league indicator	0.078*
	(0.038)
Post-trade month 3 X Inter-league indicator	0.019
	(0.071)
Post-trade month 4 X Inter-league indicator	0.066
	(0.094)
Post-trade month 1 X Pre-trade batting average	-0.462***
	(0.088)
Post-trade month 2 X Pre-trade batting average	-0.541***
	(0.094)
Post-trade month 3 X Pre-trade batting average	-0.911***
	(0.163)
Post-trade month 4 X Pre-trade batting average	-0.625*
	(0.250)
Post-trade month 1 X Inter-league indicator X Pre-trade batting average	-0.277*
	(0.127)
Post-trade month 2 X Inter-league indicator X Pre-trade batting average	-0.293*
	(0.147)
Post-trade month 3 X Inter-league indicator X Pre-trade batting average	-0.103
	(0.275)
Post-trade month 4 X Inter-league indicator X Pre-trade batting average	-0.237
	(0.372)
Observations	268,491
Number of player-seasons (fixed effects included)	706
R^2	0.0048

Note. Standard errors are clustered at the player-season level.

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Appendix D. Robustness Checks (Study 1)

I conducted a number of robustness checks to ensure that my findings presented in the paper are not the spurious results of the regression specifications and sample selection criteria. Results of my robustness checks are reported in the tables below.

1. My results are robust if I use a logistic regression, rather than an OLS regression, to predict the hitting indicator (Model 1).

2. My results are robust if I relax or strengthen the criterion I used to ensure that pre-trade batting average is granular. Specifically, I obtain qualitatively the same results if (a) I do not restrict the sample based on the number of at bats a player had prior to a trade (Model 2), (b) I only include players who had at least 50 at bats prior to a trade (Model 3), or (c) I only include players who had at least 150 at bats prior to a trade (Model 4).

3. My results are robust if I only include players who had more than 100 at bats both before and after a trade (Model 5).

4. My results are robust if I only include players who were not also a pitcher in a given season (Model 6).

5. My results are robust if I only include players who had one mid-season trade in a season (Model 7).

6. My results are robust if I control for the Batting Park Factor of the team corresponding to each observation. I control for Batting Park Factors in two ways. My first approach is to control for Batting Park Factors obtained from Baseball-Reference.com (Model 8). In this case, each team's Park Factor indicates the difference in runs scored and allowed by the team between its home and road games.[42] My second approach is to control for Batting Park Factors that are calculated specifically for hits, as my main analysis measures performance based on whether a batter successfully hit the ball each time he was at bat. FanGraphs.com provides Batting Park Factors for each type of hit (i.e., a hit that allows a batter to reach one base, two bases, three bases, or a home run). Since a hit for one base (or a *single*) is far more common than the other three types of hits, I control for FanGraphs' Park Factors for *singles* (Model 9). In this case, each team's Park Factor indicates the difference in *singles* scored and allowed by the team between its home and road games.

---

[42] http://www.baseball-reference.com/about/parkadjust.shtml. Accessed on March 16, 2015.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Post-trade indicator	0.802***	0.163***	0.163***	0.140***	0.160***
	(0.091)	(0.009)	(0.014)	(0.020)	(0.019)
Post-trade indicator X	0.334*	0.043**	0.047*	0.074*	0.073**
Inter-league indicator	(0.138)	(0.014)	(0.021)	(0.031)	(0.027)
Post-trade indicator X	-2.992***	-0.620***	-0.613***	-0.523***	-0.594***
Pre-trade batting average	(0.348)	(0.036)	(0.056)	(0.074)	(0.071)
Post-trade indicator X	-1.279*	-0.171**	-0.185*	-0.280*	-0.267*
Inter-league indicator X					
Pre-trade batting average	(0.529)	(0.055)	(0.083)	(0.117)	(0.105)
Observations	268,491	350,003	306,111	229,374	204,853
Number of player-seasons (fixed effects included)	706	1,406	895	554	474
R^2 (Pseudo R^2 for Model 1)	0.0005	0.0086	0.0054	0.0042	0.0045
Description of the robustness check	Logistic Regression	No restrictions on the number of "at bats" before a trade	More than 50 "at bats" before a trade	More than 150 "at bats" before a trade	More than 100 "at bats" both before and after a trade

Note. Standard errors are clustered at the player-season level.

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels,

Variable	Model 6	Model 7	Model 8	Model 9
Post-trade indicator	0.150***	0.148***	0.148***	0.146***
	(0.017)	(0.017)	(0.017)	(0.017)
Post-trade indicator X	0.061*	0.064*	0.064*	0.065*
Inter-league indicator	(0.026)	(0.026)	(0.026)	(0.026)
Post-trade indicator X	-0.562***	-0.556***	-0.554***	-0.548***
Pre-trade batting average	(0.067)	(0.067)	(0.067)	(0.067)
Post-trade indicator X	-0.233*	-0.248*	-0.246*	-0.252*
Inter-league indicator X				
Pre-trade batting average	(0.102)	(0.101)	(0.101)	(0.100)
Batting park factor for runs			2.82e-04	
			(2.77e-04)	
Batting park factor for singles				7.24e-04
				(5.35e-04)
Observations	263,254	267,001	268,491	268,491
Number of player-seasons (fixed effects included)	686	701	706	706
R^2 (Pseudo R^2 for Model 1)	0.0047	0.0047	0.0047	0.0047
Description of the robustness check	Only batters who were not also a pitcher	Only batters who had one mid-season trade in a given season	Controlling for batting park factor for runs	Controlling for batting park factor for singles

Note. Standard errors are clustered at the player-season level.

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Appendix E. Ancillary Analysis on Batting Park Factors (Study 1)

As explained in footnote 12 of the paper, I conducted further analysis to address the concern that batters who have a low pre-trade batting average and are traded across leagues are more likely to be traded to a team whose ballpark is favorable to batters, compared with (a) batters with a low pre-trade batting average who are traded within a league and (b) batters with a high pre-trade batting average. I used OLS regressions to predict the difference in the Batting Park Factor between a player's pre-trade team and post-trade team in a given season. The independent variables included a player's pre-trade batting average, the inter-league indicator, and their interaction (pre-trade batting average x inter-league indicator). The results of my regressions are reported in the table below. Regardless of whether the dependent variable was the difference in Batting Park Factors for *runs* (Model 1) before versus after a trade or the difference in Batting Park Factors for *singles* (Model 2), there is not a significant interaction between pre-trade batting average and the inter-league indicator (both $ps > .40$). This suggests that players with a low pre-trade batting average who are traded across leagues are not more likely than others to switch to a team with a more batter-friendly ballpark.

Variable	Model 1: Change in Batting Park Factors for runs	Model 2: Change in Batting Park Factors for singles
Pre-trade batting average	-11.542	-10.769*
	(10.729)	(5.023)
Inter-league indicator	0.009	-1.244
	(4.253)	(1.991)
Pre-trade batting average X Inter-league indicator	5.129	6.487
	(16.518)	(7.734)
Observations	706	706
Number of player-seasons	706	706
R^2	0.0101	0.0109

Note. Standard errors are clustered at the player-season level. * denotes significance at the 5% level.

Appendix F. Analysis of Alternative Mechanisms in Study 4

I used the 20-item Positive and Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988) to measure the positive affect (PA) and negative affect (NA) participants experienced after they were told that their scores would continue to be tracked (in the control condition) or that their scores on Round 2 would be tracked from zero (in the performance reset condition). Specifically, at the end of the survey, participants were instructed to recall how they felt during the break and indicate the extent to which they experienced the affective states described by the PANAS adjectives on a 5-point scale (1 = Not at all; 5 = Very much). PA scores equaled the sum of participants' ratings for the adjectives indicating positive affect (e.g., excited, alert, active, enthusiastic), and NA scores equaled the sum of participants' ratings for the adjectives indicating negative affect (e.g., distressed, hostile, scared, nervous). As shown in the table below, neither PA scores nor NA scores varied between conditions. Also, a 5,000-sample bootstrap analysis suggests that neither PA nor NA mediated the relationship between my performance reset treatment and self-assessments of pre-break performance with post-break performance.

Variable	Model 1: Positive Affect	Model 2: Negative Affect
Performance reset indicator	0.42	-0.82
	(3.13)	(1.73)
High self-assessment indicator	4.98	-3.02
	(3.10)	(1.71)
Performance reset indicator X High self-assessment indicator	-1.98	1.32
	(4.33)	(2.40)
Observations	81	81
R^2	0.05	0.05

Conditional indirect effects	Indirect effect	Bootrapped SE	95% bias corrected confidence interval
Positive affect as the mediator: Low self-assessment	0.12	1.03	[-1.61, 2.67]
Positive affect as the mediator: High self-assessment	-0.45	1.21	[-4.44, 1.05]
Negative affect as the mediator: Low self-assessment	0.11	0.93	[-1.59, 2.41]
Negative affect as the mediator: High self-assessment	-0.07	0.44	[-1.53, 0.53]

Note. Bootstrap sample size = 5,000

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively.

Appendix G. Ancillary Analysis Examining Individuals' Beliefs about Whether Performance Resets Can Improve Their Performance (Study 5)

At the end of Study 5, participants were asked to select which one of two options they believed would help them achieve a better performance on the last five games: (a) the option of resetting their statistics in a new round, (b) the option of not resetting their statistics in a new round, or (c) no differences between the two options. First, I found that 43.96 percent of participants in the low self-assessment condition believed that starting a new round would lead to better performance than not starting a new round, whereas only 29.35 percent of participants in the high self-assessment condition believed so, $\chi^2(1) = 4.21$, $p = .04$. Next, I examined whether such a difference in beliefs led people to value a performance reset differently in the high versus low self-assessment condition. I created a dummy variable, *belief in a performance reset*, to indicate whether or not participants believed that a performance reset would improve their post-break performance more than not resetting their performance statistics. When *belief in a performance reset* was included in a logistic regression to predict individuals' decision to start a new round, the positive effect of my low self-assessment manipulation on individuals' preferences for a performance reset was reduced to non-significance (from $\beta = 0.74$, $p = 0.015$ to $\beta = 0.59$, $p = 0.15$), but *belief in a performance reset* was a significant predictor ($\beta = 3.47$, $p < 0.001$). A 5,000-sample bootstrap analysis shows that the indirect effect of being in the low self-assessment condition via *belief in a performance reset* is significant (effect size $= 0.12$, $SE = 0.05$, 95% bias-corrected CI $= [0.01, 0.22]$). Altogether, as compared with those in the high self-assessment condition, participants in the low self-assessment condition were more likely to expect a performance reset to improve their performance and thus to elect to reset their statistics at a higher rate.

CHAPTER 4

GENERAL DISCUSSION AND DIRECTIONS FOR FUTURE RESEARCH

Definition and Extension of Key Constructs

This dissertation is dedicated to advancing understanding of how "fresh starts" affect motivation and individual performance. I define fresh starts as moments or events that make people feel psychologically more distant from the past than they otherwise would. I study two types of fresh starts: temporal landmarks and performance resets.

In Chapters 1 and 2, I investigate temporal landmarks, or distinct events that "stand in marked contrast to the seemingly unending stream of trivial and ordinary occurrences that happen to us everyday" (Shum, 1998, p.423). What determines whether or not a day or an event stands out as a temporal landmark? Past research suggests that days that signify the beginning of a new time period are more likely to become temporal landmarks than days in the middle of a time period (Shum, 1998; Robinson, 1986). In particular, "salient points in time that define us as individuals and members of a particular culture (i.e., occupation, religion, pastimes)" are likely to carry greater importance (Shum, 1988, p. 427). Building on these notions, Chapter 1 examines whether engagement in goal-directed activities (e.g., exercising, creating goal commitment contracts) increases at the beginning of a generic calendar cycle (e.g., the start of a week), the beginning of a new period on an academic or work calendar (e.g., the start of a semester, the first workday after a federal holiday), and the beginning of a new period in one's personal history (e.g., birthdays). Chapter 2 tests the *causal* relationship between temporal landmarks and goal initiation. Specifically, when I assign meaning and

importance to a day by highlighting that it marks the start of a new calendar or lifetime period (e.g., the first day of spring, the first day of summer break, the first day of a new zodiac cycle in China), people are more likely to initiate goal pursuit on that day.

The temporal landmarks studied in Chapters 1 and 2 are associated with either neutral or positive experiences. Temporal landmarks of negative valence (e.g., a divorce, the death of a family member) can also trigger a psychological disconnect between individuals' past and present selves. For example, people who undergo significant trauma report feeling less connected to their pre-trauma self than those at the same age without such experiences (Cantor et al., 1987); when talking about the time before receiving a cancer diagnosis, cancer patients often refer to their past self before cancer as a different person (Mathieson & Stam, 1995). An interesting question worthy of future research is whether landmark events stained by negative emotions such as grief and anger bring about feelings of a fresh start and motivate individuals to strive for self-improvement. In the short period following a negative life event, people may focus on coping with stressful experiences (Cohen & Hoberman, 1983) and may not have the cognitive and self-regulatory resources required to pursue their long-term goals. However, after people have time to process these experiences, negative landmark events may prompt them to view their lives from a different perspective and pursue aspirations that they would not have thought about otherwise. Consistent with this speculation, concurrent research (Schultz, Price, & Coulter, 2014) shows that the number of recent, significant life events people have experienced positively predicts their intentions to pursue new goals, but the

valence of those life events is not correlated with goal pursuit intentions. In other words, both positive and negative life experiences predict an uptick in goal pursuit.

In Chapter 3, I study performance resets, which I define as incidences where individuals' past performance is wiped clean such that it no longer affects how their future performance is appraised. In a field study, I examine performance resets in professional baseball. In professional baseball, being traded across leagues in the middle of a baseball season causes professional baseball players' performance after the trade to be evaluated independent of their pre-trade performance in the same season. In my laboratory experiments, I introduced performance resets in the middle of a sequence of games by tabulating some players' average performance across games from a clean slate without including their pre-reset performance. Across those studies, I focused on one-time performance resets that happen unexpectedly to performers.[43] An interesting question for future research is whether the behavioral patterns documented in Chapter 3 would emerge following recurrent performance resets. I expect that recurrent resets will still boost morale and performance among individuals who performed poorly in the previous performance period. Just as New Year's Day—a recurrent temporal landmark—motivates people who failed to achieve their resolutions in the previous year to renew their attempts, recurrent opportunities for individuals to build up their performance record from scratch may still help those with low past performance feel psychologically separated from past failures, experience higher self-efficacy, and perform better.

---

[43] Major league baseball players often do not know that they are going to be traded. One exception is that players who have been with a team for five consecutive years and have been a major league player for 10 years cannot be traded without their consent and know in advance if they are to be traded.

However, recurrent performance resets may affect individuals with strong past performance differently from one-time performance resets. Specifically, if resetting performance statistics at the beginning of a new cycle (e.g., a month, quarter) is the norm in an organization, it is unlikely that people will reduce their work commitment after a reset simply because they are discouraged by the prospect of rebuilding their performance record or because they feel that their past effort has been wasted. In particular, if each performance period is associated with a performance goal (e.g., a monthly sales goal, a quarterly earnings goal), individuals who performed well in the previous period will still experience pressure to "keep up the good work" in order to achieve strong performance again in the new performance period. In short, I expect that, as compared with one-shot performance resets, recurrent resets are less likely to damage the motivation and performance of individuals with strong past performance.

The types of fresh starts studied in this dissertation trigger a break from the past and at the same time provide an opportunity to start over. I argue that the separation from the past and the opportunity to move into the future with a clean slate are both part of fresh starts. However, the focal aspect of fresh starts (achieving closure versus starting fresh) may vary across individuals and situations. For example, individuals who devote more attention to thinking about future periods of their lives—relative to those with a weak future focus (Shipp, Edwards, & Lambert, 2009)—may be more likely to construe a transition point in their lives as an opportunity to start anew rather than as an opportunity for closure. One potential extension of this dissertation would be to examine whether engagement in goal-directed activities varies as a function of which aspect of a fresh start

is highlighted. For example, when viewing a fresh start as a chance to start over, people may be more likely to take on new challenges; however, when viewing a fresh start as closure on the past, people may be more likely to push to accomplish goals they previously initiated.

Mechanisms Driving the Effects of Temporal Landmarks and Performance Resets

The fresh start effect documented in Chapters 1 and 2 is likely to be multiply determined. Chapter 2 tests and provides support for one psychological process: temporal landmarks create a psychological disconnect between a person's current self and her past, inferior self, which motivates goal pursuit. Future research that examines additional mechanisms driving the motivating effect of temporal landmarks would be valuable. One possibility is that significant life transitions prompt people to take stock of their lives (Andreasen, 1984) and search for meaning (Alter & Hershfield, 2014), a process that may inspire them to adopt positive changes. Another possibility is that by triggering a disjuncture between individuals' past and present selves, landmark events may push individuals' present selves to feel closer to their future selves, which further leads them to behave consistently with their long-term goals (Bartels & Urminsky, 2011; Hershfield et al., 2011).

In Chapter 3, I propose that performance resets affect motivation by altering individuals' confidence in their ability to improve their performance as well as their commitment to performing well. However, performance resets may also elicit different emotional responses depending on individuals' perceptions of their past performance, which further influences their subsequent performance. If people perceive their past

performance as weak, they may feel excited and hopeful when a reset puts their bad performance record behind them; this excitement and hopefulness may boost performance (Brooks, 2014). On the other hand, if people perceive their past performance as strong, they may feel discouraged and upset when a reset prevents them from continuing to build their existing performance record and forces them to start from scratch; the resulting frustration may decrease performance (Hebb, 1955; Norman & Bobrow, 1975). Further research is needed to disentangle different potential drivers of the effects of performance resets on future performance.

The Long-term Effect of Fresh Starts on Motivation

Chapters 1 and 2 provide consistent evidence that people are more likely to initiate goal pursuit following temporal landmarks. An important open question is whether people will strive to achieve their goals more persistently and achieve better outcomes in the long run if they initiate goal pursuit after a temporal landmark than if they initiate goal pursuit on an ordinary day. Answering this question will help us understand whether managers and policymakers should encourage goal pursuit around fresh starts, particularly if repeated actions and persistent effort are needed to reach the goal in question. To draw *causal* inferences about the long-term effects of fresh starts, future research needs to randomly assign individuals to begin goal pursuit on different dates. I speculate that people assigned to begin to tackle their goals at the beginning of a new period experience higher self-efficacy, increasing their propensity to work harder and more persistently when faced with difficulties (Schmidt & DeShon, 2010; Multon, Brown, & Lent, 1991) and resulting in higher performance (Bandura & Locke, 2003).

The Effects of Anticipated Fresh Starts on Motivation

This dissertation has focused on how motivation and behaviors change *following* a temporal landmark or a performance reset. A question worthy of exploration is how behaviors change in response to the *anticipation* of a fresh start. Findings in this dissertation suggest a few possible consequences of anticipated fresh starts. First, people may believe that a better opportunity to initiate goal pursuit will arise following an impending landmark and thus delay launching their plans until after the landmark. Consistent with this speculation, ancillary analysis of data tracking stickK.com commitment contracts (Study 3 of Chapter 1) reveals that 20% of stickK users choose to start their commitment contracts on a future date (as opposed to now), and those users are significantly more likely to commit themselves in advance to starting a contract at the beginning of a new calendar period (e.g., Monday, the first workday after a holiday).

Second, people may use upcoming fresh starts as self-imposed deadlines and strive to bring *ongoing* goals to closure (e.g., finish reading a book, paying off a debt) by these deadlines because they want to avoid bringing unfinished businesses into a new mental accounting period. In a similar vein, when people have accumulated satisfying progress (e.g., abstaining from smoking for 11 months, obtaining perfect attendance at work for 27 days), the impending arrival of a new performance period may motivate them to keep up the good work, as they would feel remorseful about ruining a perfect record at the end of a period.

Third, anticipated fresh starts might liberate people to make goal-incongruent choices if they anticipate wiping the slate clean after an upcoming fresh start (Zhang,

Fishbach, & Dhar, 2007). In particular, people may reduce effort devoted to performance improvement if they are discouraged by their performance so far (Casas-Arce, & Martinez-Jerez, 2009; Hannan, Krishnan, & Newman, 2008), hoping that they will have a chance to start over in the next performance period. Future work is needed to explore different ways in which people vary their engagement in goal-related activities in anticipation of fresh starts.

The Creation of Self-initiated Fresh Starts

Across three chapters, I have primarily examined situations where external events (e.g., a birthday, a trade between sports teams) psychologically separate people from the previous calendar or performance period. Can individuals internally create a break from the past? Ample anecdotes suggest that when people want to leave their past behind them and start anew, they are able to construct fresh starts themselves. Self-generated fresh starts can be as significant as moving to a new residence (e.g., Kaufman, 2013; the film *Rebound,* 1999) and starting a journey around the world (e.g., the film *Eat, Pray, Love*, 2010) or as minor as clearing out cars or closets (Senise, 2014; Biziou, 2015), removing tattoos (Jackson, 2013), and getting a new hairstyle (Biziou, 2015). Consistent with these examples, Study 5 in Chapter 3 shows that when people view their past performance as poor (rather than strong), they are more likely to elect to reset their performance statistics and create a new performance period. It would be valuable for future work to examine (i) factors that motivate people to enact a fresh start and (ii) the effectiveness of different techniques or rituals people may leverage to regulate the psychological connectedness between their past and future selves.

Implications of Fresh Starts for Behaviors in Domains Other Than Goal Pursuit and Performance Improvement

In addition to goal pursuit and performance management, fresh starts may also affect behaviors in other important domains. One area that is worth exploring is ethical decision making. Past research has documented a slippery slope effect whereby minor violations of ethical rules can cause moral disengagement and increasingly unethical behaviors (Welsh, Ordóñez, Snyder, & Christian, 2014). Techniques that help people to (a) reflect on the temptations they have given into and (b) start over with a clean slate may reduce their propensity to slide down the slippery slope. Notwithstanding, findings in Chapter 3 suggest one caveat: introducing a fresh start may also lead individuals who previously did not cheat to feel licensed to act unethically.

Integrating fresh starts with the literature on escalation of commitment is another interesting future avenue. One common explanation for why people engage in escalation behavior is that they feel personally responsible for their decision failures and experience the need to justify their initial decisions (Staw, 1976; Sleesman, Conlon, McNamara, & Miles, 2012). Framing a situation in relationship to a fresh start may psychologically separate individuals from their past selves who made the bad decision, reduce self-justification needs, and thus alleviate their tendency to escalate commitment.

Further, fresh starts may have implications for interpersonal relationships. For example, people may be more likely to forgive another person who previously hurt them after they experience a fresh start (e.g., moving to a new city, switching to a new job) because the fresh start—which may be unrelated to the transgression in question—helps

people to leave the unpleasant past behind. Also, it is plausible that people make more lenient decisions if someone who committed a wrongdoing recently had a fresh start (e.g., a milestone birthday, the birth of a baby)—particularly a fresh start imbued with the concept of repentance (e.g., Rosh Hashanah for Jews, the Sacrament of Confession for Catholics), because people may believe that the transgressor is now ready to start anew. In summary, future research should examine how fresh starts affect people's beliefs about others' ability to change their behaviors and further influence interpersonal judgments.

References

Alter, A. L., & Hershfield, H. E. (2014). People search for meaning when they approach a new decade in chronological age. *Proceedings of the National Academy of Sciences of the United States of America, 111*(48), 17066-17070. doi: 10.1073/pnas.1415086111

Andreasen, A. R. (1984). Life status changes and changes in consumer preferences and satisfaction. *Journal of Consumer Research, 11*(3), 784-794. doi: 10.1086/209014

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87-99. doi: 10.1037/0021-9010.88.1.87

Bartels, D. M., & Urminsky, O. (2011). On intertemporal selfishness: How the perceived instability of identity underlies impatient consumption. *Journal of Consumer Research, 38*(1), 182-198. doi: 10.1086/658339

Biziou. B. (2015). Chinese year of the sheep: Rituals for a fresh start! The Huffington Post. Retrieved on April 20, 2015 from http://www.huffingtonpost.com/barbara-biziou/rituals-for-a-fresh-start_b_6694676.html.

Brooks, A. W. (2014). Get excited: Reappraising pre-performance anxiety as excitement. *Journal of Experimental Psychology-General, 143*(3), 1144-1158. doi: 10.1037/A0035325

Cantor, N., Norem, J. K., Niedenthal, P. M., Langston, C. A., & Brower, A. M. (1987). Life tasks, self-concept ideals, and cognitive strategies in a life transition. *Journal of Personality and Social Psychology, 53*(6), 1178-1191. doi: 10.1037/0022-3514.53.6.1178

Casas-Arce, P., & Martinez-Jerez, F. A. (2009). Relative performance compensation, contests, and dynamic incentives. *Management Science, 55*(8), 1306-1320. doi: 10.1287/mnsc.1090.1021

Cohen, S., & Hoberman, H. M. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology, 13*(2), 99-125. doi: 10.1111/j.1559-1816.1983.tb02325.x

Hannan, R. L., Krishnan, R., & Newman, A. H. (2008). The effects of disseminating relative performance feedback in tournament and individual performance compensation plans. *Accounting Review, 83*(4), 893-913. doi: 10.2308/accr.2008.83.4.893

Hebb, D. O. (1955). Drives and the cns (conceptual nervous system). Psychological Review, 62(4), 243-254. doi: 10.1037/H0041823

Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L., Carstensen, L. L., & Bailenson, J. N. (2011). Increasing saving behavior through age-progressed renderings of the future self. *Journal of Marketing Research, 48*, S23-S37.

Jackson. C. (2013). A fresh start. *The Story*. Retrieved on April 20, 2015 from http://www.thestory.org/stories/2013-06/fresh-start.

Kaufman, J. (2013). Relocation therapy. *The New York Times*. Retrieved on January 4, 2013 from http://www.nytimes.com/2013/01/06/realestate/relocation-therapy.html?pagewanted=all&_r=0.

Mathieson, C. M., & Stam, H. J. (1995). Renegotiating identity - cancer narratives. *Sociology of Health and Illness, 17*(3), 283-306. doi: 10.1111/1467-9566.Ep10933316

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes - a meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30-38. doi: 10.1037/0022-0167.38.1.30

Norman, D. A., & Bobrow, D. G. (1975). Data-limited and resource-limited processes. *Cognitive Psychology, 7*(1), 44-64. doi: 10.1016/0010-0285(75)90004-3

Robinson, J. A. (1986). Temporal reference systems and autobiographical memory. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 159-188). Cambridge, England: Cambridge University Press.

Schmidt, A. M., & DeShon, R. P. (2010). The moderating effects of performance ambiguity on the relationship between self-efficacy and performance. *Journal of Applied Psychology, 95*(3), 572-581. doi: 10.1037/A0018289

Schultz, A., Price, L., & Coulter, R. (2014). Embracing a "Fresh Start": How Consumers Engage to Change Their Lives. Working Paper, University of Arizona, Tucson, AZ.

Senise. M. (2014). Cleaned out my car, cleaning out my life. *The Huffington Post*. Retrieved on April 20, 2015 from http://www.huffingtonpost.com/maria-senise/cleaned-out-my-car-cleani_b_5692047.html.

Shipp, A. J., Edwards, J. R., & Lambert, L. S. (2009). Conceptualization and measurement of temporal focus: The subjective experience of the past, present,

and future. *Organizational Behavior and Human Decision Processes, 110*(1), 1-22. doi: 10.1016/j.obhdp.2009.05.001

Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory processes. *Psychological Bulletin, 124*(3), 423-442. doi: 10.1037//0033-2909.124.3.423

Sleesman, D. J., Conlon, D. E., McNamara, G., & Miles, J. E. (2012). Cleaning up the big muddy: A meta-analytic review of the determinants of escalation of commitment. *Academy of Management Journal, 55*(3), 541-562. doi: 10.5465/amj.2010.0696

Staw, B. M. (1976). Knee-deep in big muddy - study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance, 16*(1), 27-44. doi: 10.1016/0030-5073(76)90005-2

Welsh, D. T., Ellis, A. P. J., Christian, M. S., & Mai, K. M. (2014). Building a self-regulatory model of sleep deprivation and deception: The role of caffeine and social influence. *Journal of Applied Psychology, 99*(6), 1268-1277. doi: 10.1037/A0036202

Zhang, Y., Fishbach, A., & Dhar, R. (2007). When thinking beats doing: The role of optimistic expectations in goal-based choice. *Journal of Consumer Research, 34*, 567-578.